

Louisiana State University LSU Digital Commons

LSU Master's Theses

Graduate School

2003

A parametrization approach for solving the Hamilton-Jacobi-Equation and application to the A2 Toda lattice

Mohammad Dikko Aliyu

Louisiana State University and Agricultural and Mechanical College, maliyu1@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses



Part of the [Applied Mathematics Commons](#)

Recommended Citation

Aliyu, Mohammad Dikko, "A parametrization approach for solving the Hamilton-Jacobi-Equation and application to the A2 Toda lattice" (2003). *LSU Master's Theses*. 2539.

https://digitalcommons.lsu.edu/gradschool_theses/2539

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

A PARAMETRIZATION APPROACH FOR SOLVING THE HAMILTON-JACOBI-EQUATION AND APPLICATION TO THE \mathcal{A}_2 TODA LATTICE

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The Department of Mathematics

by

Mohammad Dikko Aliyu
B.Eng. E.E. A. B. U. Zaria, Nigeria, June, 1988
M.S. E.E. Louisiana State University, May, 2001
PhD, L.S.U., May, 2002
May, 2003

Dedicated to my late father
May God have mercy on him

ACKNOWLEDGEMENTS

I wish to express my profound gratitude to my family for their patience and prayers during my graduate studies. I am highly indebted to Prof. L. Smolinsky for the guidance and motivation that he has provided throughout the period of my graduate studies, and my appreciation goes to Profs. G. Ferreyra and M. Tom for participating in my thesis committee.

Lastly, I wish to express my thanks to Louisiana State University for the support it has provided during my graduate studies.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
ABSTRACT	v
CHAPTER	
1 INTRODUCTION TO HAMILTON-JACOBI THEORY	1
1.1 Introduction	1
1.2 The Hamiltonian Formulation of Mechanics	2
1.3 The Transformation Generating Function	8
1.4 The Hamilton-Jacobi Equation	10
1.4.1 The Hamilton-Jacobi Equation for the Time-Independent Hamiltonian	11
1.5 Separation of Variables in the Hamilton-Jacobi Equation	12
1.5.1 Geometry of the Hamilton-Jacobi Equation	13
1.6 The Theory of Nonlinear Lattices	15
1.6.1 The \mathcal{G}_2 -Periodic Toda Lattice	18
2 THE METHOD OF CHARACTERISTICS FOR FIRST-ORDER PARTIAL DIFFERENTIAL EQUATIONS AND ITS RELATIONSHIP TO HAMILTON'S CANONICAL EQUATIONS	19
2.1 Introduction	19
2.2 The Method of Characteristics	19
2.3 Characteristic Equations for the General First-Order Equation	21
2.3.1 Characteristic Equations for the Hamilton-Jacobi Equation	22
2.4 Characteristics for the General First-Order Equation in Two Variables	23
2.5 Geometric View-Point of the General First-Order Equation	25
2.5.1 The Characteristic Curves of the PDE	27
2.5.2 Initial Conditions	28
2.5.3 Legendre Transform and Hopf-Lax Formula	29
3 Solving the Hamilton-Jacobi Equation and Application to the Toda Lattice	32
3.1 Introduction	32
3.2 Solving the Hamilton-Jacobi Equation	32
3.2.1 Solution of the Hamilton-Jacobi equation for the \mathcal{A}_2 Toda system	37
REFERENCES	42
VITA	43

ABSTRACT

Hamilton-Jacobi (HJ)-theory is an extension of Lagrangian mechanics and concerns itself with a directed search for a coordinate transformation in which the equations of motion can be easily integrated. The equations of motion of a given mechanical system can often be simplified considerably by a suitable transformation of variables such that all the new position and momentum coordinates are constants. A particular type of transformation is chosen in such a way that the new equations of motion retain the same form as in the former coordinates; such a transformation is called *canonical* or *contact* and can greatly simplify the solution to the equations of motion. Hamilton (1838) has developed the method for obtaining the desired transformation equations using what is today known as *Hamilton's principle*. It turns out that the required transformation can be obtained by finding a smooth function S called a *generating function* or *Hamilton's principal function*, which satisfies a certain nonlinear first-order partial-differential equation (PDE) also known as the *Hamilton-Jacobi equation* (HJE).

Unfortunately, the HJE being nonlinear, is very difficult to solve; and thus, it might appear that little practical advantage has been gained in the application of the HJ-theory. Nonetheless, under certain conditions, and when the Hamiltonian is independent of time, it is possible to separate the variables in the HJE, and the solution can then always be reduced to quadratures. Thus, the HJE becomes a useful computational tool only when such a separation of variables can be achieved.

However, in this thesis we develop another approach for solving the HJE for a large class of Hamiltonian systems in which the variables may not be separable and/or the Hamiltonian is not time-independent. We apply the approach to a class of integrable Hamiltonian systems known as the Toda lattice. Computational results are presented to show the usefulness of the method.

CHAPTER 1

INTRODUCTION TO HAMILTON-JACOBI THEORY

1.1 Introduction

Hamiltonian mechanics is a transformation theory that is an off-shoot of Lagrangian mechanics. It concerns itself with a directed search for coordinate transformations which exhibit specific advantages for certain types of problems; notably in celestial and quantum mechanics. As such, the Hamiltonian approach to the analysis of a dynamical system, as it stands, does not represent an overwhelming development over the Lagrangian method. One ends up with practically the same number of equations as the Lagrangian approach. Thus, the real advantage of this approach lies in a technique whereby the transformed equations of motion in terms of a new set of position and momentum variables are easily integrated for specific problems, and the deeper insight it provides into the formal structure of mechanics. The equal status accorded to coordinates and momenta as independent variables provides a new representation and greater freedom in selecting more relevant coordinate systems for different types of problems.

In this Thesis, we study Lagrangian systems from the Hamiltonian standpoint. We shall consider natural mechanical systems for which the kinetic energy is a positive-definite quadratic form of the generalized velocities, and the Lagrangian function is the difference between the kinetic energy and the potential energy. Furthermore, as will be reviewed shortly, it will be shown that, the Hamiltonian transformation of the equations of motion of a mechanical system always lead to the Hamilton-Jacobi (HJ) partial differential equation (PDE) which is a first-order nonlinear PDE that must be solved in order to obtain the required transformation generating function. However, as is well-known from the theory of PDEs, the HJE being nonlinear is very difficult to solve except for the case when the Hamiltonian function is such that the variables are separable. It is therefore our objective in this thesis to present a method for solving the HJE for a class of Hamiltonian systems that may not admit a separation of variables. We shall then apply the method to a class of Hamiltonian systems known as the Toda lattice which are integrable [16].

The rest of the thesis is organized as follows. In the remainder of this chapter we shall review the Hamiltonian formulation of the equations of motion of a mechanical system that will culminate in the Hamilton-Jacobi equation (HJE). We also discuss various aspects of the HJE including separation of variables and its geometry. We then review the theory of Toda lattices as a class of Hamiltonian systems that are integrable. We discuss in particular the nonperiodic and the period \mathcal{A}_2 -Toda lattices, and the \mathcal{G}_2 -periodic Toda lattices.

In chapter 2, we discuss the method of characteristics for solving first-order nonlinear equations in two or more independent variables. It is shown that the characteristic equations that result in the transformation of the HJE are exactly the Hamilton's canonical equations, and in this respect, we show that the method actually owes its origin from Hamilton-Jacobi

theory. We discuss both the case of two independent variables and n independent variables. Furthermore, we also discuss a geometric view-point to the method. Finally, at the end of the chapter, we discuss Legendre transforms and the Hopf-Lax formula for solving the HJE.

In chapter 3, we discuss our new approach for solving the HJE. We derive sufficient conditions for the solvability of the HJE and estimate bounds on the solution and its derivatives. We then apply the method to solve the HJE for the two-particle nonperiodic \mathcal{A}_2 Toda lattice. We also integrate the equations of motion to get the system trajectories.

Finally, we give conclusion in chapter 4.

Notation: The notation is fairly standard except where otherwise stated. Moreover, \mathfrak{R} , \mathfrak{R}^n will denote respectively, the real line and the n -dimensional real vector space, $t \in \mathfrak{R}$ will denote the time parameter. Let M^n , N^n, \dots denote Riemannian manifolds with dimension n , which are compact. Let $TM = \bigcup_{x \in M} T_x M$, $T^*M = \bigcup_{x \in M} T_x^* M$ respectively denote the tangent and cotangent bundles of M with dimensions $2n$. Moreover, π_M and π_M^* will denote the natural projections $TM \rightarrow M$ and $T^*M \rightarrow M$ respectively. $SO(n, M)$ and $SM(n, M)$ will denote the special orthogonal group and the set of symmetric matrices over M .

A $C^\infty(M)$ vector-field is a mapping $f : M \rightarrow TM$ such that $\pi \circ f = I_M$ (the identity on M), and f has continuously differentiable partial derivatives of arbitrary order. A vector field f also defines a differential equation (or a dynamic system) $\dot{x}(t) = f(x)$, $x \in M$, $x(t_0) = x_0$. The flow (or integral curve) of the differential equation $\phi(t, x_0)$, $t \in \mathfrak{R}$, is the unique solution of the differential equation for any arbitrary initial condition x_0 over an open interval $I \subset \mathfrak{R}$. The flow of a differential equation will also be referred as the trajectory of the system and will be denoted by $x(t, x_0)$ or $x(t)$ when the initial condition is immaterial. We shall also assume throughout this paper that the vector fields are complete, and hence the domain of the flow extends over $(-\infty, \infty)$. Furthermore, an equilibrium point of the vector field f or the differential equation defined by it, is a point \bar{x} such that $f(\bar{x}) = 0$ or $\phi(t, \bar{x}) = \bar{x} \forall t \in \mathfrak{R}$. An *invariant set* for the system $\dot{x}(t) = f(x)$, is any set \mathcal{A} such that, for any $x_0 \in \mathcal{A}$, $\Rightarrow \phi(t, x_0) \in \mathcal{A}$ for all $t \in (-\infty, \infty)$. A differential k -form ω_x^k , $k = 1, 2, \dots$, at a point $x \in M$ is an exterior product of covectors from $T_x^* M$ to \mathfrak{R} i.e. $\omega_x^k : T_x^* M \times \dots \times T_x^* M$ (k copies) $\rightarrow \mathfrak{R}$, which is a k -linear skew-symmetric function of k -vectors on $T_x^* M$. The space of all smooth k -forms on M is denoted by $\Omega^k(M)$ which are the smooth C^∞ sections of the vector bundle $\Lambda^k(T^*M)$. The \mathcal{F} -derivative (Frèchet derivative) of a real-valued function $U : \mathfrak{R}^n \rightarrow \mathfrak{R}$ is defined as any ϱ such that $\lim_{v \rightarrow 0} \frac{1}{\|v\|} [U(x + v) - U(x) - \langle \varrho, v \rangle] = 0$, for any $v \in \mathfrak{R}^n$. For a smooth function $f : \mathfrak{R}^n \rightarrow \mathfrak{R}$, $f_x = \frac{\partial f}{\partial x} = (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})$. Further, let $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty : M \rightarrow \mathfrak{R}$ denote respectively, 1, 2, and ∞ norms on M , where $\|v(q)\|_1 = \sum_{i=1}^n |v_i(q)|$, $\|v(q)\|_2 = \sqrt{\sum_{i=1}^n |v_i(q)|^2}$ and $\|v(q)\|_\infty = \max_i \{|v_i(q)| : i = 1, \dots, n\}$ for any vector $v : M_q \rightarrow T_q M$. Also, if $f : [0, 1] \rightarrow \mathfrak{R}$, then $\|f(s)\|_{\mathcal{L}_p} = \left(\int_0^1 |f(s)|^p \right)^{\frac{1}{p}}$, $0 < p < \infty$.

1.2 The Hamiltonian Formulation of Mechanics

To review the approach, let the configuration space of the system be defined by a smooth n -dimensional Riemannian manifold M . If (φ, U) is a coordinate chart, we write $\varphi = q =$

(q_1, \dots, q_n) for the local coordinates and $\dot{q}_i = \frac{\partial}{\partial q_i}$ in the tangent bundle $TM|_U = TU$. We shall be considering *natural mechanical systems* which are defined as follows.

Definition 1.2.1 *A Lagrangian mechanical system on a Riemannian manifold is called natural if the Lagrangian function $L : TU \times \mathbb{R} \rightarrow \mathbb{R}$, with $\mathcal{U} \subset M$ open, is equal to the difference between the kinetic energy and the potential energy of the system as*

$$L(q, \dot{q}, t) = T(q, \dot{q}, t) - V(q, t), \quad (1.1)$$

where $T : \mathcal{U} \rightarrow \mathbb{R}$ is the kinetic energy which is given by the quadratic form

$$T = \frac{1}{2} \langle v, v \rangle, \quad v \in T_q U$$

and $V : M \times \mathbb{R} \rightarrow \mathbb{R}$ is the potential energy of the system (which may be independent of time).

For natural mechanical systems, the kinetic energy is a positive-definite symmetric quadratic form of the generalized velocities,

$$T(q, \dot{q}, t) = \frac{1}{2} \dot{q}^T \Psi(q, t) \dot{q}. \quad (1.2)$$

It is further known from Lagrangian mechanics and as can be derived using the D'Alembert's *principle of virtual work* or Hamilton's *principle of least action* [5, 9, 10] (Theorem 1.2.2), that the equations of motion of a *holonomic conservative*¹ mechanical system satisfy Lagrange's equations of motion given by

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, \dots, n. \quad (1.3)$$

Then the above equation (1.3) may always be written in the form

$$\ddot{q} = g(q, \dot{q}, t), \quad (1.4)$$

for some function $g : \mathcal{U} \rightarrow \mathbb{R}^n$.

On the other hand, in the Hamiltonian formulation, we choose to replace all the \dot{q}_i by independent coordinates, p_i , in such a way that

$$p_i := \frac{\partial L}{\partial \dot{q}_i} \quad i = 1, \dots, n, \quad (1.5)$$

If we let

$$p_i = h(q, \dot{q}), \quad i = 1, \dots, n, \quad (1.6)$$

then the Jacobian of h with respect to \dot{q} , using (1.1), (1.2) and (1.5), is given by $\Psi(q)$ which is positive definite, and hence equation (1.5) can be inverted to yield

$$\dot{q}_i = f_i(q_1, \dots, q_n, p_1, \dots, p_n, t), \quad i = 1, \dots, n. \quad (1.7)$$

¹Holonomic if the constraints on the system are expressible as equality constraints. Conservative if there exists a time-dependent potential.

for some continuous functions f_1, \dots, f_n . In this framework, the coordinates $q = (q_1, q_2, \dots, q_n)^T$ are referred to as the *generalized coordinates* and $p = (p_1, p_2, \dots, p_n)^T$ are the *generalized momenta*. Together, these variables form a new system of coordinates for the system known as the *phase space* of the system. If (U, φ) where $\varphi = (q_1, q_2, \dots, q_n)$ is a chart on M , then since $p_i : TU \rightarrow \mathfrak{R}$, they are elements of T^*U , and together with the q_i 's form a system of $2n$ local coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$, where $p_i(q) \in T_q^*M$, $i = 1, \dots, n$, for the phase-space.

We now define the Hamiltonian function of the system $H : T^*M \times \mathfrak{R} \rightarrow \mathfrak{R}$ as the *Legendre transform*² of the Lagrangian function with respect to \dot{q} by

$$H(q, p, t) = p^T \dot{q} - L(q, \dot{q}, t). \quad (1.8)$$

Consider now the differential of H with respect to q , p and t as

$$dH = \left(\frac{\partial H}{\partial p} \right)^T dp + \left(\frac{\partial H}{\partial q} \right)^T dq + \frac{\partial H}{\partial t} dt. \quad (1.9)$$

The above expression must be equal to the total differential of $H = p\dot{q} - L$ for $p = \frac{\partial L}{\partial \dot{q}}$:

$$dH = \dot{q}^T dp - \left(\frac{\partial L}{\partial q} \right)^T dq - \left(\frac{\partial L}{\partial t} \right)^T dt \quad (1.10)$$

Thus, in view of the independent nature of the coordinates, we obtain a set of three relationships:

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \frac{\partial L}{\partial q} = -\frac{\partial H}{\partial q}, \quad \text{and} \quad \frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t}.$$

Finally applying Lagrange's equation (1.3) together with (1.5) and the preceeding results, one obtains the expression for \dot{p} . Since we used Lagrange's equation, $\dot{q} = \frac{dq}{dt}$ and $\dot{p} = \frac{dp}{dt}$. The resulting Hamiltonian canonical equations of motion are then given by

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}(q, p, t), \quad (1.11)$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q}(q, p, t), \quad (1.12)$$

Thus, we have proven the following theorem.

Theorem 1.2.1 *The system of Lagrange's equations (1.3) is equivalent to the system of $2n$ first-order Hamilton's equations (1.11), (1.12).*

In addition, for time-independent conservative systems, $H(q, p)$ has a simple physical interpretation. From (1.8) and using (1.5), we have

$$\begin{aligned} H(q, p, t) &= p^T \dot{q} - L(q, \dot{q}, t) = \dot{q}^T \frac{\partial L}{\partial \dot{q}} - (T(q, \dot{q}, t) - U(q, t)) = \dot{q}^T \frac{\partial T}{\partial \dot{q}} - T(q, \dot{q}, t) + U(q, t) \\ &= 2T(q, \dot{q}, t) - T(q, \dot{q}, t) + U(q, t) = T(q, \dot{q}, t) + U(q, t), \end{aligned} \quad (1.13)$$

²To be defined later, see also [5]

i.e., the total energy of the system. This completes the Hamiltonian formulation of the equations of motion, and can be seen as an off-shoot of the Lagrangian formulation. It can be seen that, while the Lagrangian formulation involves n second-order equations, the Hamiltonian description sets up a system of $2n$ first-order equations in terms of the $2n$ variables p and q . This remarkably new system of coordinates gives new insight and physical meaning to the equations. However, the system of Lagrange's equations and Hamilton's equations are completely equivalent and dual to one another.

Furthermore, because of the symmetry of Hamilton's equations (1.11), (1.12) and the even dimension of the system, a new structure emerges on the phase space T^*M of the system. This structure is defined by a nondegenerate closed differential 2-form which in the above local coordinates is defined as :

$$\omega^2 = dp \wedge dq = \sum_{i=1}^n dp_i \wedge dq_i. \quad (1.14)$$

Thus, the pair (T^*M, ω^2) form a *symplectic manifold*, and together with a C^r Hamiltonian function $H : T^*M \rightarrow \mathfrak{R}$ define a Hamiltonian mechanical system. With this notation we have the following representation of a Hamiltonian system.

Definition 1.2.2 *Let (T^*M, ω^2) be a symplectic manifold and $H : T^*M \rightarrow \mathfrak{R}$ the Hamiltonian function. Then the vector field X_H determined by the condition*

$$\omega^2(X_H, Y) = dH(Y) \quad (1.15)$$

*for all vector fields Y , is called the Hamiltonian vector field with energy function H . Call the tuple (T^*M, ω^2, X_H) a Hamiltonian system.*

Remark 1.2.1 *It is important to note that, the nondegeneracy of ω^2 guarantees that X_H exists, and is a C^{r-1} vector field. Moreover, on a connected symplectic manifold, any two Hamiltonians for the same vector field X_H have the same differential (1.15), so differ by a constant only.*

We also have the following proposition [1].

Proposition 1.2.1 *Let $(q_1, \dots, q_n, p_1, \dots, p_n)$ be canonical coordinates so that ω^2 is given by (1.14). Then, in these coordinates*

$$X_H = \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n} \right) = J \cdot \nabla H$$

where $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$. Thus, $(q(t), p(t))$ is an integral curve of X_H if and only if Hamilton's equations (1.11), (1.12) hold.

Now suppose that a transformation of coordinates is introduced $q_i \rightarrow Q_i$, $p_i \rightarrow P_i$, $i = 1, \dots, n$ such that every Hamiltonian function transforms as $H(q_1, \dots, q_n, p_1, \dots, p_n, t) \rightarrow$

$K(Q_1, \dots, Q_n, P_1, \dots, P_n, t)$ in such a way that the new equations of motion retain the same form as in the former coordinates, i.e.,

$$\frac{dQ}{dt} = \frac{\partial K}{\partial p}(Q, P, t) \quad (1.16)$$

$$\frac{dP}{dt} = -\frac{\partial K}{\partial q}(Q, P, t) \quad (1.17)$$

Such a transformation is called *canonical* or *contact* and can greatly simplify the solution to the equation of motion, especially if Q, P are selected such that $K(., ., .)$ is a constant independent of Q and P . Should this happen, then Q and P will also be constants and the solution to the equations of motion are immediately at hand (given the transformation). We simply transform back to the original coordinates; under the assumption that the transformation is univalent and invertible. It therefore follows from this that:

1. The inverse of a canonical transformation is a canonical transformation;
2. The product of two canonical transformations is itself a canonical transformation;
3. A canonical transformation must preserve the 2-form $\omega^2 = dp \wedge dq$.

Hamilton (1838) has developed a method for obtaining the desired transformation equations using what is today known as *Hamilton's principle* which states as follows.

Definition 1.2.3 Let $\gamma = \{(t, q) : q = q(t), t_0 \leq t \leq t_1\}$ be a curve in the (t, q) plane. Define the functional $\Phi(\gamma)$ (which we assume to be differentiable) by

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(q(\tau), \dot{q}(\tau)) d\tau.$$

Then, the curve γ is an extremal of the functional $\Phi(.)$ if $\delta\Phi(\gamma) = 0$ or $d\Phi(\gamma) = 0 \forall t \in [t_0, t_1]$, where δ is the variational operator.

Theorem 1.2.2 (Hamilton's principle of least action) [5, 9, 10, 12]. The motion of a mechanical system with Lagrangian function $L(., ., .)$, coincide with the extremals of the functional $\Phi(\gamma)$.

Accordingly, define the Lagrangian function of the system $L : \mathcal{U} \rightarrow \mathfrak{R}$ as the Legendre transform [5] of the Hamiltonian function by

$$L(q, \dot{q}, t) = p^T \dot{q} - H(q, p, t). \quad (1.18)$$

Then, in the new coordinates, the new Lagrangian function is

$$\bar{L}(Q, \dot{Q}, t) = P^T \dot{Q} - K(Q, P, t). \quad (1.19)$$

Since both $L(., ., .)$ and $\bar{L}(., ., .)$ are conserved, each must separately satisfy Hamilton's principle. However, $L(., ., .)$ and $\bar{L}(., ., .)$ need not be equal in order to satisfy the above requirement. Indeed we can write

$$L(q, \dot{q}, t) = \bar{L}(Q, \dot{Q}, t) + \frac{dS}{dt}(q, p, Q, P, t) \quad (1.20)$$

for some arbitrary function $S : \mathcal{X} \times \mathcal{X} \rightarrow \mathfrak{R}$, where $\mathcal{X} \subset T^*M$ is open ([10], page 286). Since dS is an exact differential,

$$\delta \left[\int_{t_0}^{t_1} \frac{dS}{dt}(q, p, Q, P, t) dt \right] = S(q, p, Q, P, t)|_{t_0}^{t_1} = 0. \quad (1.21)$$

Now applying Hamilton's principle to the time integral of both sides of equation (1.20), we get

$$\delta \left[\int_{t_0}^{t_1} L(q, \dot{q}, t) dt \right] = \delta \left[\int_{t_0}^{t_1} \bar{L}(Q, \dot{Q}, t) dt \right] + \delta \left[\int_{t_0}^{t_1} \frac{dS}{dt}(q, p, Q, P, t) dt \right] = 0; \quad (1.22)$$

and Therefore by (1.21),

$$\delta \left[\int_{t_0}^{t_1} \bar{L}(Q, \dot{Q}, t) dt \right] = 0. \quad (1.23)$$

Thus, to guarantee that a given change of coordinates, say,

$$q_i = \phi_i(Q, P, t) \quad (1.24)$$

$$p_i = \psi_i(Q, P, t) \quad (1.25)$$

is canonical, from (1.18), (1.19) and (1.20), it is enough that

$$p^T \dot{q} - H = P^T \dot{Q} - K + \frac{dS}{dt}. \quad (1.26)$$

This condition is also required [10]. Hence the above equation is equivalent to

$$p^T dq - P dQ = (H - K)(q, p, Q, P, t) dt + dS(q, p, Q, P, t). \quad (1.27)$$

Requiring upon the expression on the left side to be an exact differential!

We can verify that the presence of $S(.)$ in (1.20) does not alter the canonical character of the Hamiltonian equations of motion. In this regard, applying Hamilton's principle to the RHS of (1.20), we have from (1.23), the Euler-Lagrange equation (1.3), and the argument following it

$$\frac{dQ}{dt} = \frac{\partial K}{\partial p}(Q, P, t) \quad (1.28)$$

$$\frac{dP}{dt} = -\frac{\partial K}{\partial q}(Q, P, t), \quad (1.29)$$

hence, preserving the canonical nature of the equations of motion.

1.3 The Transformation Generating Function

A given Hamiltonian system can often be simplified considerably by a suitable transformation of variables such that all the new position and momentum coordinates (Q_i, P_i) are constants. A particular type of transformation is discussed in this section.

We have already seen that an arbitrary generating function, S , does no injury to our equations of motion. The next step is to show that, first, if such a function is known, then the transformation we so desperately seek follows directly. Secondly, that the function can be obtained by solving a certain partial differential equation.

The generating function S relates the old to the new coordinates via the equation

$$S = \int (L - \bar{L}) dt = f(q, p, Q, P, t) \quad (1.30)$$

Thus, S is a function of $4n + 1$ variables, and hence no more than four independent sets of relationships among the dependent coordinates can exist. Two such relationships expressing the old sets of coordinates in terms of the new set are given by equations (1.24), (1.25). Hence only two independent sets of relationships among the coordinates remain for defining S and no more than two of the four sets of coordinates may be involved. Therefore, there are four possibilities:

$$S_1 = f_1(q, Q, t); \quad S_2 = f_2(q, P, t) \quad (1.31)$$

$$S_3 = f_3(p, Q, t); \quad S_4 = f_4(p, P, t) \quad (1.32)$$

Any one of the above four types of generating functions may be selected, and a transformation obtained from it. For example, if we consider the generating function S_1 , taking its differential, we have

$$dS_1 = \sum_{i=1}^n \frac{\partial S_1}{\partial q_i} dq_i + \sum_{i=1}^n \frac{\partial S_1}{\partial Q_i} dQ_i + \frac{\partial S_1}{\partial t} dt. \quad (1.33)$$

Again, taking the differential as defined by (1.27), we have

$$dS_1 = \sum_{i=1}^n p_i dq_i - \sum_{i=1}^n P_i dQ_i + (K - H) dt. \quad (1.34)$$

Finally, using the independence of coordinates, we equate coefficients, and obtain the desired transformation equations

$$\left. \begin{aligned} p_i &= \frac{\partial S_1}{\partial q_i}(q, Q, t) \\ P_i &= -\frac{\partial S_1}{\partial Q_i}(q, Q, t) \\ K - H &= \frac{\partial S_1}{\partial t}(q, Q, t) \end{aligned} \right\}, \quad i = 1, \dots, n. \quad (1.35)$$

Similar derivation can be applied to the remaining three types of generating functions, except that we can also apply Legendre transformation. Thus, for the generating functions $S_2(., ., .)$, $S_3(., ., .)$ and $S_4(., ., .)$, we have

$$\left. \begin{aligned} p_i &= \frac{\partial S_2}{\partial q_i}(q, P, t) \\ Q_i &= -\frac{\partial S_2}{\partial P_i}(q, P, t) \\ K - H &= \frac{\partial S_2}{\partial t}(q, P, t) \end{aligned} \right\}, \quad i = 1, \dots, n, \quad (1.36)$$

$$\left. \begin{aligned} q_i &= -\frac{\partial S_3}{\partial p_i}(p, Q, t) \\ P_i &= -\frac{\partial S_3}{\partial Q_i}(p, Q, t) \\ K - H &= \frac{\partial S_3}{\partial t}(p, Q, t) \end{aligned} \right\}, i = 1, \dots, n, \quad (1.37)$$

$$\left. \begin{aligned} q_i &= -\frac{\partial S_4}{\partial p_i}(p, P, t) \\ Q_i &= \frac{\partial S_4}{\partial P_i}(p, P, t) \\ K - H &= \frac{\partial S_4}{\partial t}(p, P, t) \end{aligned} \right\}, i = 1, \dots, n, \quad (1.38)$$

respectively.

It should however be remarked that, the great generality of the canonical transformations expressed in the arbitrariness of the choice of generating functions often has the consequence that the distinct meaning of the generalized coordinates and momenta is completely lost. For example, consider the generating function $S = S_1(q, Q) = q^T Q$. Then it follows from the foregoing, that

$$\left. \begin{aligned} p_i &= \frac{\partial S_1}{\partial q_i} = Q_i \\ P_i &= -\frac{\partial S_1}{\partial Q_i} = -q_i \\ K &= \frac{\partial S_1}{\partial t} = H(-P, Q, t) \end{aligned} \right\}, i = 1, \dots, n, \quad (1.39)$$

hold, that is, with the exception of a sign change, the momenta and coordinates have been interchanged.

The simplest type of canonical transformation is that for which $Q(q, p, t)$ does not depend on p but only on q and possibly t . In that case, the coordinates are transformed only among themselves; this special kind of canonical transformation is known as a *point transformation*. Whereas a generic canonical transformation might destroy the distinct meaning of the coordinates and momenta (coordinates describe the position; momenta together with the coordinates, the velocity), this meaning is preserved for point transformations. Consider for instance the transformation $Q = f(q, t)$ of the coordinates among each other may be derived from the generating function

$$S = S_2(q, P, t) = f^T(q, t)P \quad (1.40)$$

and

$$\left. \begin{aligned} p_i &= \frac{\partial S_2}{\partial q_i} = \left(\frac{\partial f}{\partial q}\right)^T P, \\ Q_i &= \frac{\partial S_2}{\partial P_i} = f(q, t), \\ K &= H + \frac{\partial f^T(q, t)}{\partial t} P \end{aligned} \right\}, i = 1, \dots, n. \quad (1.41)$$

The above equations (1.41) provide the transformation of the momenta and the Hamiltonian for a given point transformation based on $f(q, t)$.

Next, given a transformation $(q, p) \rightarrow (Q, P)$, it is generally not immediately clear whether such a transformation is canonical or not. It is therefore desirable to have a criterion

which may be used to determine whether if such a transformation is canonical. In this case, one tool is to use Poisson brackets. For given functions $f(q, p, t)$, $g(q, p, t)$, both in C^1 , their Poisson bracket is defined as

$$[f, g] = \left(\frac{\partial f}{\partial q} \right)^T \left(\frac{\partial g}{\partial p} \right) - \left(\frac{\partial f}{\partial p} \right)^T \left(\frac{\partial g}{\partial q} \right). \quad (1.42)$$

It can then be shown that the transformation T is canonical if and only if (iff):

$$[Q_i, Q_k] = 0, \quad [P_i, P_k] = 0, \quad [P_i, Q_k] = \delta_{ik}, \quad i, k = 1, 2, \dots, n \quad (1.43)$$

are satisfied, where δ_{ik} is the Kronecker delta.

1.4 The Hamilton-Jacobi Equation

In this section, we turn our attention to the last missing link in the Hamiltonian transformation theory; an approach for determining the transformation generating function, S . There is only one equation available for this purpose

$$H(q, p, t) + \frac{\partial S}{\partial t} = K(P, Q, t). \quad (1.44)$$

However, there are two unknown functions in this equation: S and K . Thus, the best we can do is to assume a solution for one and then solve for the other. In this regard, suppose we arbitrarily introduce the condition that K is to be *identically zero*? Under this condition, \dot{Q} and \dot{P} vanish; resulting in $Q = \alpha$, and $P = \beta$, constants. The inverse transformation then yields the motion $q(\alpha, \beta, t)$, $p(\alpha, \beta, t)$ in terms of these constants of integration, α and β .

Consider now generating functions of the first type. Having forced a solution on K , we must now solve the partial differential equation (PDE)

$$\mathbf{H}(\mathbf{q}, \frac{\partial \mathbf{S}}{\partial \mathbf{q}}, \mathbf{t}) + \frac{\partial \mathbf{S}}{\partial \mathbf{t}} = 0 \quad (1.45)$$

for S , where $\frac{\partial S}{\partial \mathbf{q}} = (\frac{\partial S}{\partial q_1}, \dots, \frac{\partial S}{\partial q_n})^T$. This equation is known as the *Hamilton-Jacobi Equation (HJE)*, and was improved and modified by Jacobi in 1838. For a given function $H(q, p, t)$, this is a first-order PDE in the unknown function $S(q, \alpha, t)$ which is customarily called *Hamilton's principal function*. We need a solution for this equation which depends on n arbitrary constants $\alpha_1, \alpha_2, \dots, \alpha_n$ in such a way that the Jacobian determinant of $\frac{\partial S}{\partial q_i}$ with respect to (wrt) the α_j satisfies

$$\left| \frac{\partial^2 S}{\partial q_i \partial \alpha_j} \right| \neq 0. \quad (1.46)$$

The above condition excludes the possibility in which one of the n constants α_j is additive; that is, one must have

$$S(q, \alpha, t) \neq \bar{S}(q, \alpha_1, \alpha_2, \dots, \alpha_{n-1}, t) + \alpha_n \quad (1.47)$$

A solution $S(q, \alpha, t)$ satisfying (1.46) is called a 'complete solution' of the HJE (1.45), and solving the HJE is equivalent to finding the solutions of the equations of motion (1.11),

(1.12). Conversely, the solution of (1.45) is nothing more than a solution of the equations (1.11), (1.12) using the method of characteristics [7]. However, it is generally not simpler to solve (1.45) instead of (1.11), (1.12).

If a complete solution $S(q, \alpha, t)$ of (1.45) is known, then one has

$$\frac{\partial S}{\partial q_i} = p_i, \quad i = 1, \dots, n \quad (1.48)$$

$$\frac{\partial S}{\partial \alpha_i} = -\beta_i, \quad i = 1, \dots, n \quad (1.49)$$

Since the condition (1.46) is satisfied, the second algebraic equation above may be solved for q and the first solved for $p(\alpha, \beta, t)$. One thus has a canonical transformation from (α, β) to (q, p) . And it follows from the definition of canonical transformation that the inverse transformation $\alpha = \alpha(q, p, t)$, $\beta = \beta(q, p, t)$ also is canonical. The following theorem is by Hamilton-Jacobi [1].

Theorem 1.4.1 *Let \widetilde{X}_H be the Hamiltonian vector-field on $T^*M = \mathfrak{R} \times T^*M$ corresponding to the system (1.11), (1.12) and let $S : T^*M \rightarrow \mathfrak{R}$. Then the following conditions are equivalent:*

(i) *for every curve $c(t)$ in $\widetilde{T^*M}$ satisfying*

$$c'(t) = T \pi_M^* (X_{H_t}(dS_t(c(t))))$$

the curve $t \mapsto dS_t(c(t))$ is an integral curve of X_H ;

(ii) *S satisfies the HJE (1.45).*

1.4.1 The Hamilton-Jacobi Equation for the Time-Independent Hamiltonian

The preceding section has laid down a systematic approach to the solution of the equations of motion via a transformation theory that culminates in the HJE. However, implementation of the above procedure is difficult; because the prospects of success are limited by the inadequate state of the mathematical art in regard to the solution of PDE - especially nonlinear PDEs. At present, the only technique of general utility is the method of *separation of variables*. If the Hamiltonian is explicitly a function of time, then separation of variables is not readily achieved for the HJE. However, if on the other hand, the Hamiltonian is not explicitly a function of time or is independent of time, which arises in many dynamical systems of practical interest, then the HJE separates easily. The solution to (1.45) can then be formulated in the form

$$S(q, \alpha, t) = -ht + W(q, \alpha) \quad (1.50)$$

with $h = h(\alpha)$. Consequently, the use of (1.50) in (1.45) yields the following PDE in W :

$$H(q, \frac{\partial W}{\partial q}) = h, \quad (1.51)$$

where h is the energy constant (if the kinetic energy of the system is homogeneous quadratic, the constant equals the total energy, E). Moreover, since W does not involve time, the new and the old Hamiltonians are equal, and it follows that $K = h$. The function W , known as *Hamilton's characteristic function*, thus generates a canonical transformation in which all the new coordinates are cyclic. Further, one may choose $h = \alpha_n$ for example, so that

$$W = W(q, \alpha_1, \dots, \alpha_{n-1}, h) \quad (1.52)$$

depends on $n - 1$ additional arbitrary constants besides h . Noting that the Jacobian determinant of S wrt the n arbitrary coordinates, and the n constants $\alpha_1, \dots, \alpha_{n-1}, h$ may not vanish, then from (1.48), (1.49) and (1.50), we have the following system

$$\left. \begin{aligned} \frac{\partial W}{\partial \alpha_i} &= -\beta_i, \quad i = 1, 2, \dots, n-1, \\ \frac{\partial W}{\partial h} &= t - \beta_n, \\ \frac{\partial W}{\partial q} &= p. \end{aligned} \right\}, \quad (1.53)$$

where the term $t - \beta_n$ in the preceeding equation follows directly from the fact that the system is autonomous. The above system of equations may be solved for $n - 1$ components of q , say, for q_1, q_2, \dots, q_{n-1} resulting in

$$\left. \begin{aligned} q_1 &= q_1(\alpha_1, \alpha_2, \dots, \alpha_n, h, \beta_1, \beta_2, \dots, \beta_{n-1}, q_n), \\ q_2 &= q_2(\alpha_1, \alpha_2, \dots, \alpha_n, h, \beta_1, \beta_2, \dots, \beta_{n-1}, q_n), \\ \vdots &= \vdots \\ q_{n-1} &= q_{n-1}(\alpha_1, \alpha_2, \dots, \alpha_n, h, \beta_1, \beta_2, \dots, \beta_{n-1}, q_n), \end{aligned} \right\} \quad (1.54)$$

where the time t is replaced as the parameter q_n . These equations are then the solution for the system, and describe its trajectories. The following theorem is also by Hamilton-Jacobi [1] for the time-independent case.

Theorem 1.4.2 *Let X_H be the Hamiltonian vector-field on T^*M corresponding to the system (1.11), (1.12) and let $S : T^*M \rightarrow \mathfrak{R}$. Then, the following conditions are equivalent:*

(i) *for every curve $c(t)$ in T^*M satisfying*

$$c'(t) = T \pi_M^* (X_H(dS(c(t))))$$

the curve $t \mapsto dS(c(t))$ is an integral curve of X_H ;

(ii) *S satisfies the HJE (1.51).*

1.5 Separation of Variables in the Hamilton-Jacobi Equation

In this section, we briefly take up the issue again of the practical advantages of the application of the HJ-theory in solving the equations of motion of a given mechanical system. It might appear from the preceding section that little practical advantage has been gained

in solving a first-order nonlinear PDE, which is notoriously complicated to solve, instead of a system of $2n$ ODEs. Nonetheless, under certain conditions, and when the Hamiltonian is independent of time, it is possible to separate the variables in the HJE, and the solution can then always be reduced to quadratures. Thus, the HJE becomes a useful computational tool only when such a separation of variables can be achieved.

We consider the case when the Hamiltonian is independent of time and the HJE is of the form (1.51). The variables q_i occurring in this equation are said to be separable if a solution of the form

$$W(q, \alpha) = \sum_i W_i(q_i, \alpha_1, \dots, \alpha_n) \quad (1.55)$$

splits the HJE into n equations of the form

$$H_i(q_i, \frac{\partial W_i}{\partial q_i}, \alpha_1, \dots, \alpha_n) = \alpha_i \quad (1.56)$$

Each of the equations (1.56) involves only one of the coordinates q_i and the corresponding partial derivative of W_i wrt q_i . They are therefore a set of n first-order ODEs of a particularly simple form which can be solved by quadratures.

There is no simple criterion for determining when the HJE is separable. For some problems eg. the three-body problem, it is impossible to separate the variables, while for others it is transparently easy. Fortunately, a great majority of systems of current interest in quantum mechanics and atomic physics, are of the latter class. It should also be emphasized that, the question of whether the HJE is separable depends on the system of generalized coordinates employed. Indeed, the one-body central force problem is separable in polar coordinates, but not in cartesian coordinates.

1.5.1 Geometry of the Hamilton-Jacobi Equation

In this subsection, we discuss briefly some additional geometry connected with the Hamilton-Jacobi equation (1.45). Most of the discussion in this section is from [1, 5, 13]. For each $\tilde{q} = (q, t) \in \tilde{M} := M \times \mathbb{R}$, $dS(\tilde{q}) \in T\tilde{M}$. As \tilde{q} varies in \tilde{M} , the set $\{\mathbf{d}S(q) \mid \tilde{q} \in \tilde{M}\}$ defines a submanifold of $T^*\tilde{M}$ that in terms of coordinates is given by $p_j = \tilde{p}_j = \frac{\partial S}{\partial q_j}$, $j = 1, \dots, n$ and $p_{n+1} = \tilde{p}_{n+1} = \frac{\partial S}{\partial t}$. We call this submanifold the *range* or graph of $\mathbf{d}S$, i.e.,

$$\text{graph}(\mathbf{d}S) := \{(\tilde{q}, \tilde{p} = \frac{\partial S}{\partial q_j}), j = 1, \dots, n+1) \mid \tilde{q} \in \tilde{M}\}.$$

Furthermore, recall that [5] the configuration space T^*M of the mechanical system together with the nondegenerate 2-form $\omega = dp \wedge dq \in \Omega^2(M)$ is a symplectic manifold. Thus, the extended configuration space $T^*\tilde{M}$ together with the 2-form $\tilde{\omega} = d\tilde{p} \wedge d\tilde{q}$ form a symplectic manifold of dimension $2n+2$. A submanifold $\tilde{N} \subset T^*\tilde{M}$ whose dimension is n and $\omega|_{\tilde{N}} \equiv 0$ is called a *Lagrangian submanifold*. Therefore, since the restriction of the 2-form $\tilde{\omega}$ to the $\text{graph}(\mathbf{d}S)$ is given by

$$\tilde{\omega}|_{\text{graph}(\mathbf{d}S)} = d\frac{\partial S}{\partial \tilde{q}} \wedge d\tilde{q} = \frac{\partial^2 S}{\partial \tilde{q}^2} d\tilde{q} \wedge d\tilde{q} = 0,$$

and the dimension of $\text{graph}(\mathbf{d}S)$ is $n+1$, then $\text{graph}(\mathbf{d}S) \subset T^*\tilde{M}$ is a Lagrangian submanifold. What is important about such a submanifold is that, the projection $\tilde{\pi}_M^* : \text{graph}(\mathbf{d}S) \rightarrow \tilde{M}$ is a diffeomorphism, and even more, the converse holds: If $\Lambda \subset T^*\tilde{M}$ is a Lagrangian submanifold such that the projection on \tilde{M} is a diffeomorphism in a neighborhood of a point $\lambda \in \Lambda$, then in some neighborhood of λ , then in some neighborhood of λ we can write $\Lambda = \text{graph}(\mathbf{d}\varphi)$ for some function φ . To show this, notice that because the projection is a diffeomorphism, Λ is given (around λ) as a submanifold of the form $(\tilde{q}_j, \rho_j(\tilde{q}))$. The condition for Λ to be Lagrangian requires that on Λ ,

$$\begin{aligned} \sum_{j=1}^{n+1} d\tilde{q}_j \wedge d\tilde{p} &= 0, \\ \Downarrow \\ \sum_{j=1}^{n+1} d\tilde{q}_j \wedge d\tilde{\rho}(\tilde{q}) &= 0, \\ \Downarrow \\ \frac{\partial \rho_j}{\partial \tilde{q}_k} - \frac{\partial \rho_k}{\partial \tilde{q}_j} &= 0; \end{aligned}$$

thus, there exists a φ such that $\rho_j = \frac{\partial \varphi}{\partial \tilde{q}_j}$, which is the same as $\Lambda = \text{graph}(\mathbf{d}S)$. The conclusion of these remarks is that Lagrangian submanifolds of $T^*\tilde{M}$ are natural generalizations of the graphs of differentials of functions on \tilde{M} . Note however that, Lagrangian submanifolds are defined even if the projection on \tilde{M} is not a diffeomorphism.

From the point of view of Lagrangian submanifolds, the graph of the differential of a solution of the HJE is a Lagrangian submanifold of $T^*\tilde{M}$ that is contained in the surface $\tilde{\mathcal{S}}_{\tilde{H}} \subset T^*\tilde{M}$, where $\tilde{H}(\tilde{q}, \tilde{p}) = H(t, q, p) + \tilde{p}_{n+1}$ is the corresponding Hamiltonian function which is possibly time-dependent. This point of view allows one to include solutions that are singular in the usual context. This is not the only benefit: We also get more insight into the HJE. The tangent space to $\tilde{\mathcal{S}}_{\tilde{H}}$ has dimension 1 less than the dimension of the manifold $T^*\tilde{M}$, and it is given by the set of vectors X such that $(d\tilde{p}_{n+1} + \mathbf{d}H)(X) = 0$. If a vector Y is in the symplectic orthogonal [1] of $T_{(\tilde{q}, \tilde{p})}(\tilde{\mathcal{S}}_{\tilde{H}})$, i.e.,

$$\sum_{j=1}^{n+1} d\tilde{q}_j \wedge d\tilde{p}(X, Y) = 0$$

for all $x \in T_{(\tilde{q}, \tilde{p})}(\tilde{\mathcal{S}}_{\tilde{H}})$, then, Y is a multiple of the vector field

$$X_{\tilde{H}} = \frac{\partial}{\partial t} - \frac{\partial H}{\partial t} \frac{\partial}{\partial p} + X_H$$

evaluated at (\tilde{q}, \tilde{p}) . Moreover, the integral curves of $X_{\tilde{H}}$ projected to $(q, p) \in T^*M$ are solutions of the HJE for H .

The key observation that links Hamilton's equations and the HJE is that the vector field $X_{\tilde{H}}$ is tangent to $\tilde{\mathcal{S}}$ and to any Lagrangian submanifold contained in $\tilde{\mathcal{S}}$, which is equivalent to saying that a solution for the HJE is either disjoint from a Lagrangian submanifold or completely contained in it. This observation gives a way of constructing a solution to the HJE starting from an initial condition at $t = t_0$. Namely, start with a Lagrangian submanifold $\Lambda_0 \subset T^*M$ and embed it in $T^*\tilde{M}$ at $t = t_0$ using

$$(q, p) \mapsto (q, t = t_0, p, p_{n+1} = -H(q, p, t_0)).$$

The result is an isotropic submanifold $\tilde{\Lambda}_0 \subset T^*\tilde{M}$ (i.e., a submanifold on which the canonical 2-form vanishes [1]). Now take all the integral curves of $X_{\tilde{H}}$ whose initial conditions lie in $\tilde{\Lambda}_0$. The collection of these curves spans a manifold Λ whose dimension is 1 higher than the dimension of $\tilde{\Lambda}_0$. Since $X_{\tilde{H}}$ is tangent to $\tilde{\mathcal{S}}_{\tilde{H}}$ and $\Lambda_0 \subset \tilde{\mathcal{S}}_{\tilde{H}}$, we get $\Lambda_t \subset \tilde{\mathcal{S}}_{\tilde{H}}, \forall t$, and hence $\Lambda \in \tilde{\mathcal{S}}_{\tilde{H}}$, where $\Lambda_t = \Phi_t(\tilde{\Lambda}_0)$ and Φ_t is the flow of $X_{\tilde{H}}$. Moreover, since Φ_t is a canonical map, it leaves $T^*\tilde{M}$ invariant and therefore takes an isotropic submanifold into an isotropic one; in particular $\Lambda_t \subset T^*\tilde{M}$ is isotropic. Upon further analysis [13], it can be shown that Λ is also an isotropic submanifold of $T^*\tilde{M}$. Furthermore, since its dimension is half the dimension of $T^*\tilde{M}$, it follows that it is Lagrangian and contained in $\tilde{\mathcal{S}}_{\tilde{H}}$. Thus, it is a solution of the HJE with initial condition Λ_0 at $t = t_0$.

1.6 The Theory of Nonlinear Lattices

Historically, the the exact treatment of oscillations in nonlinear lattices became serious in the early 1950's when Fermi, Pasta, and Ulam (FPU) numerically studied the problem of energy partition. Fermi et al. wanted to verify by numerical experiment if there is energy flow between the modes of linear-lattice systems when nonlinear interactions are introduced - he wanted to verify what is called the *equipartition of energy* in statistical mechanics. However, to their dissapointment, only a little energy partition occurred, and the state of the systems was found to return periodically to the initial state.

Later, Ford and co-workers [16] showed that by using pertubation and by numerical calculation, though resonance generally enhances energy sharing, it has no intimate connection to a periodic phenomenon, and that nonlinear lattices have rather stable-motion (periodic, when the energy is not too high) or pulses (also known as solitons), which he called the *nonlinear normal modes*. This fact also indicate that there will be some nonlinear lattice which admits rigorous periodic waves, and certain pulses (lattice solitons) will be stable there. This remarkable property led to the finding of an integrable 1-dimensional lattice with exponential interaction also known as the *Toda lattice*.

The Toda lattice as a Hamiltonian system describes the motion of n particles moving in a straight line, with “exponential interaction” between them. Mathematically, it is equivalent to a problem in which a single particle moves in \mathbb{R}^n . Let the positions of the particles at time t (in \mathbb{R}) be $q_1(t), \dots, q_n(t)$, respectively. We assume that each particle has mass 1. The momentum of the i -th particle at time t is therefore $p_i = \dot{q}_i$. The Hamiltonian function for the *finite (or non-periodic)* lattice is defined to by

$$H(q, p) = \frac{1}{2} \sum_{j=1}^n p_j^2 + \sum_{j=1}^{n-1} e^{2(q_j - q_{j+1})}. \quad (1.57)$$

Thus the canonical equations for the system are given by

$$\left. \begin{aligned} \frac{dq_j}{dt} &= p_j \quad j = 1, \dots, n, \\ \frac{dp_1}{dt} &= -2e^{2(q_1 - q_2)}, \\ \frac{dp_j}{dt} &= -2e^{2(q_j - q_{j+1})} + 2e^{2(q_{j-1} - q_j)}, \quad j = 2, \dots, n-1, \\ \frac{dp_n}{dt} &= 2e^{2(q_{n-1} - q_n)}. \end{aligned} \right\} \quad (1.58)$$

It may be assumed in addition that $\sum_{j=1}^n q_j = \sum_{j=1}^n p_j = 0$, and the coordinates q_1, \dots, q_n can be chosen so that this condition is satisfied. While for the periodic lattice in which the

first particle interacts with the last, the Hamiltonian function is defined by

$$\tilde{H}(q, p) = \frac{1}{2} \sum_{j=1}^n p_j^2 + \sum_{j=1}^{n-1} e^{2(q_j - q_{j+1})} + e^{2(q_n - q_1)}. \quad (1.59)$$

We may also consider the infinite lattice, in which there are infinitely many particles.

Nonlinear lattices can provide models for nonlinear phenomena such as wave propagation in nerve systems, chemical reactions, certain ecological systems and a host of electrical and mechanical systems. For example, it is easily shown that a linear lattice is equivalent to a ladder network composed of capacitors \mathbf{C} and inductors \mathbf{L} , while a one-dimensional nonlinear lattice is equivalent to a ladder circuit with nonlinear \mathbf{L} or \mathbf{C} . To show this, let I_n denote the current, Q_n the charge on the capacitor, Φ_n the flux in the inductance, and write the equations for the circuit as

$$\left. \begin{aligned} \frac{dQ_n}{dt} &= I_n - I_{n-1}, \\ \frac{d\Phi_n}{dt} &= V_n - V_{n+1} \end{aligned} \right\} \quad (1.60)$$

Now assume that the inductors and capacitors are nonlinear in such a way that

$$\begin{aligned} Q_n &= C v_0 \ln(1 + V_n/v_0) \\ \Phi_n &= L i_0 \ln(1 + I_n/i_0) \end{aligned}$$

where (C, v_0, L, i_0) are constants. Then equations (1.60) give

$$\begin{aligned} \frac{dQ_n}{dt} &= i_0 \left(e^{\frac{\Phi_{n-1}}{L i_0}} - e^{\frac{\Phi_n}{L i_0}} \right) \\ \frac{d\Phi_n}{dt} &= v_0 \left(e^{\frac{Q_{n-1}}{C v_0}} - e^{\frac{Q_n}{C v_0}} \right) \end{aligned}$$

which are in form of a the lattice with exponential interaction (or Toda system).

Stimulated by Ford's numerical work which revealed the likely integrability of the Toda lattice, Henon and Flaschka [16] independently showed the integrability of the Toda lattice analytically, and began an analytical survey of the lattice. At the same time, the inverse scattering method of solving the initial value problem for the Kortoweg-de Vries equation (KdV) had been firmly formulated by Lax [16], and this method was applied to the infinite lattice to derive a solution for the using matrix formalism which led to a simplification of the equations of motion. To introduce this formalism, define the following matrices

$$L = \begin{pmatrix} p_1 & Q_{1,2} & 0 & \cdots & 0 & 0 \\ Q_{1,2} & p_2 & Q_{2,3} & \cdots & 0 & 0 \\ 0 & Q_{2,3} & p_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p_{n-1} & Q_{n-1,n} \\ 0 & 0 & 0 & \cdots & Q_{n-1,n} & p_n \end{pmatrix} \quad (1.61)$$

$$M = \begin{pmatrix} 0 & Q_{1,2} & 0 & \cdots & 0 & 0 \\ -Q_{1,2} & 0 & Q_{2,3} & \cdots & 0 & 0 \\ 0 & -Q_{2,3} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & Q_{n-1,n} \\ 0 & 0 & 0 & \cdots & -Q_{n-1,n} & 0 \end{pmatrix} \quad (1.62)$$

where $Q_{i,j} = e^{(q_i - q_j)}$. We then have the following proposition [11].

Proposition 1.6.1 *The Hamiltonian system for the non-periodic Toda lattice (1.58)-(1.58) is equivalent to the Lax equation $\dot{L} = [L, M]$, where the function L, M take values in $sl(n, \mathbb{R})$ and $[\cdot, \cdot]$ is the Lie bracket operation in $sl(n, \mathbb{R})$.*

Using the above matrix formalism, the solution of the Toda system (1.58) can be derived [11, 16].

Theorem 1.6.1 *The solution of the Hamiltonian system for the Toda lattice is given by $L(t) = Ad(\exp tV)_I^{-1}V$, where $V = L(0)$ and I represents the identity matrix.*

The solution can be explicitly written in the case of $n = 2$. Letting $q_1 = -q$, $q_2 = q$, $p_1 = -p$ and $p_2 = p$, we have

$$L = \begin{pmatrix} p & Q \\ Q & -p \end{pmatrix}, \quad M = \begin{pmatrix} 0 & Q \\ -Q & 0 \end{pmatrix},$$

where $Q = c^{-2q}$. The solution of $\dot{L} = [L, M]$ with

$$L(0) = \begin{pmatrix} 0 & v \\ v & 0 \end{pmatrix},$$

is

$$L(t) = Ad \left[\exp t \begin{pmatrix} 0 & v \\ v & 0 \end{pmatrix} \right]_I^{-1} \begin{pmatrix} 0 & v \\ v & 0 \end{pmatrix}.$$

Now

$$\exp t \begin{pmatrix} 0 & v \\ v & 0 \end{pmatrix} = \begin{pmatrix} \cosh tv & \sinh tv \\ \sinh tv & \cosh tv \end{pmatrix},$$

and hence,

$$\left[\exp t \begin{pmatrix} 0 & v \\ v & 0 \end{pmatrix} \right]_I^{-1} = \frac{1}{\sqrt{\sinh^2 tv + \cosh^2 tv}} \begin{pmatrix} \cosh tv & \sinh tv \\ \sinh tv & \cosh tv \end{pmatrix}.$$

Therefore,

$$L(t) = \frac{v}{\sinh^2 tv + \cosh^2 tv} \begin{pmatrix} -2 \sinh tv \cosh tv & 1 \\ 1 & 2 \sinh tv \cosh tv \end{pmatrix}.$$

Which means that

$$p(t) = -v \frac{\sinh 2tv}{\cosh 2tv}, \quad Q(t) = \frac{v}{\cosh 2tv}.$$

Furthermore, if we recall that $Q(t) = e^{-2q(t)}$, it follows that

$$\begin{aligned} q(t) &= -\frac{1}{2} \log \left(\frac{v}{\cosh 2tv} \right) \\ &= -\frac{1}{2} \log v + \frac{1}{2} \log \cosh 2vt. \end{aligned} \quad (1.63)$$

1.6.1 The \mathcal{G}_2 -Periodic Toda Lattice

In the study of the generalized periodic Toda lattice, Bogoyavlensky [6] showed that various models of the Toda lattices which admit the $[L, M]$ -Lax representation, correspond to certain simple Lie-algebras which he called the $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ and \mathcal{G}_2 periodic Toda systems. In particular, the \mathcal{G}_2 is a two-particle system and corresponds to the Lie algebra g_2 which is 14-dimensional, and has been studied extensively in the literature [2, 3, 14]. The Hamiltonian for the g_2 system is given by

$$H(q, p) = \frac{1}{2}(p_1^2 + p_2^2) + e^{(1/\sqrt{3})q_1} + e^{-(\sqrt{3}/2)q_1 + (1/2)q_2} + e^{-q_2}, \quad (1.64)$$

and the Lax equation corresponding to this system is given by $dA/dt = [A, B]$, where

$$\begin{aligned} A(t) &= a_1(t)(X_{-\beta_3} + X_{\beta_3}) + a_2(t)(X_{-\gamma_1} + X_{\gamma_1}) + a_3(t)(s^{-1}X_{-\gamma_3} + sX_{\gamma_3}) + b_1(t)H_1 \\ &\quad + b_2H_2 \\ B(t) &= a_1(t)(X_{-\beta_3} - X_{\beta_3}) + a_2(t)(X_{-\gamma_1} - X_{\gamma_1}) + a_3(t)(s^{-1}X_{-\gamma_3} - sX_{\gamma_3}) \end{aligned}$$

where s is a parameter, the $\beta_i, i = 1, 2, 3$ and the $\gamma_j, j = 1, 2, 3$ are the short and long roots of the g_2 root system, while $X_{(-)}$ are the corresponding Chevalley basis vectors. Using the following change of coordinates [14]:

$$\begin{aligned} a_1(t) &= \frac{1}{2\sqrt{6}}e^{(1/2\sqrt{3})q_1(t)}, \quad a_2(t) = \frac{1}{2\sqrt{2}}e^{-(\sqrt{3}/4)q_1(t) + (1/4)q_2(t)}, \quad a_3(t) = \frac{1}{2\sqrt{2}}e^{-(1/2)q_2(t)}, \\ b_1(t) &= \frac{-1}{2\sqrt{3}}p_1(t) + \frac{1}{4}p_2(t), \quad b_2(t) = \frac{1}{2\sqrt{3}}p_1(t), \end{aligned}$$

we can represent the g_2 lattice as

$$\dot{a}_1 = a_1 b_2, \quad \dot{a}_2 = a_2(b_1 - b_2), \quad \dot{a}_3 = a_3(-2b_1 - b_2), \quad (1.65)$$

$$\dot{b}_1 = 2(a_1^2 - a_2^2 + a_3^2), \quad \dot{b}_2 = -4a_1^2 + 2a_2^2, \quad (1.66)$$

$$H = \frac{1}{2}\langle A(t), A(t) \rangle = 8(3a_1^2 + a_2^2 + a_3^2 + a_3^2 + b_1^2 + b_1 b_2 + b_2^2). \quad (1.67)$$

Here, the coordinate $a_2(t)$ may be regarded as superfluous, and can be eliminated using the fact that $4a_1^3 a_2^2 a_3 = c$ (a constant) of the motion.

CHAPTER 2

THE METHOD OF CHARACTERISTICS FOR FIRST-ORDER PARTIAL DIFFERENTIAL EQUATIONS AND ITS RELATIONSHIP TO HAMILTON'S CANONICAL EQUATIONS

2.1 Introduction

In this chapter, we present the well-known *method of characteristics* for solving first-order PDEs. It is by far the most generally recognized method for handling first-order nonlinear equations in n independent variables. It involves converting the PDE into an appropriate system of first-order ordinary differential equations (ODE), which are in turn solved together to obtain the solution of the original PDE. It will be seen during the development that, the Hamilton's canonical equations are nothing but the characteristic equations of the Hamilton-Jacobi equation; and thus, solving the canonical equations is equivalent to solving the PDE and vice-versa.

We shall present an exhaustive discussion of the method because it forms the backbone of our approach to the HJE. Moreover, the presentation will follow closely those from Fritz-Johns [8] and Evans [7].

2.2 The Method of Characteristics

We begin with a motivational discussion of the method by considering quasi-linear equations of which will belong the auxiliary equations of our transformation approach to the HJE (to be presented in the next chapter), and then we consider the general first-order nonlinear equations.

The general first-order equation for a function $u = u(x, y, \dots, z)$ is of the form

$$f(x, y, \dots, z, u, u_x, \dots, u_z) = 0. \quad (2.1)$$

The HJE and many first-order PDEs in classical and continuum mechanics, calculus of variations, and geometric optics are of the above type. A simpler case of the above equation is the two-variable *quasi-linear equation* :

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u) \quad (2.2)$$

in two independent variables x, y . The function $u(x, y)$ is represented by a surface $z = u(x, y)$ called an *integral surface* which corresponds to a solution of the PDE. The functions $a(x, y, z)$, $b(x, y, z)$ and $c(x, y, z)$ define a field of vectors in the xyz -space, while $(u_x, u_y, -1)$ is the normal to the surface $z = u(x, y)$.

We associate to the field of characteristic directions (a, b, c) a family of *characteristic curves* that are tangent to these directions. Along any characteristic curve $(x(t), y(t), z(t))$, where t is parameter, the following system of ODEs must be satisfied:

$$\frac{dx}{dt} = a(x, y, z), \quad \frac{dy}{dt} = b(x, y, z), \quad \frac{dz}{dt} = c(x, y, z) \quad (2.3)$$

If a surface $S : z = u(x, y)$ is a union of characteristic curves, then S is an integral surface; for then through any point P of S , there passes a characteristic curve Γ contained in S .

Next, we consider the *Cauchy problem* for the quasi-linear equation (2.2). It is desired to find a definite method for finding solutions of the PDE given an initial data on the problem. A simple way of selecting a particular candidate solution $u(x, y)$ out of an infinite set of solutions, consists in prescribing a curve Γ in xyz -space which is to be contained in the integral surface $z = u(x, y)$. Without any loss of generality, we can represent Γ parametrically by

$$x = f(s), \quad y = g(s), \quad z = h(s), \quad (2.4)$$

and we seeking for a solution $u(x, y)$ such that

$$h(s) = u(f(s), g(s)), \quad \forall s. \quad (2.5)$$

The above problem is the Cauchy problem for (2.2). Our first aim is to derive conditions for a local solution to (2.2) in the vicinity of $x_0 = f(s_0)$, $y_0 = g(s_0)$. For the sake of the analysis, we may take y to be the time parameter, and thus $u(x, 0) = h(x)$ as our initial data. This parameterizes Γ as

$$x = s, \quad y = g(s), \quad z = h(s).$$

Let now the functions $f(s), g(s), h(s) \in C^1$ in the neighborhood of some point P_0 be parameterized by s_0 , i.e.,

$$P_0 = (x_0, y_0, z_0) = (f(s_0), g(s_0), h(s_0)). \quad (2.6)$$

Assume also the coefficients $a(x, y, z), b(x, y, z), c(x, y, z) \in C^1$ near P_0 . Then we can describe Γ near P_0 by the solution

$$x = X(s, t), \quad y = Y(s, t), \quad z = Z(s, t) \quad (2.7)$$

of the characteristic equations (2.3) which reduces to $f(s), g(s), h(s)$ at $t = 0$. Therefore, the functions X, Y, Z must satisfy

$$X_t = a(X, Y, Z), \quad Y_t = b(X, Y, Z), \quad Z_t = c(X, Y, Z) \quad (2.8)$$

identically in s, t and also satisfy the initial conditions

$$X(s, 0) = f(s), \quad Y(s, 0) = g(s), \quad Z(s, 0) = h(s). \quad (2.9)$$

By the theorem on existence and uniqueness of solutions to systems of ODEs, it follows that there exists unique set of functions $X(s, t), Y(s, t), Z(s, t)$ of class C^1 satisfying (2.8), (2.9)

for (s, t) near $(s_0, 0)$. Further, if we can solve equation (2.7) for s, t in terms of x, y , say $s = S(x, y)$ and $t = T(x, y)$, then z can be expressed as

$$z = u(x, y) = Z(S(x, y), T(x, y)), \quad (2.10)$$

which represents an integral surface Σ . By (2.6), (2.9), $x_0 = X(s_0, 0)$, $y_0 = Y(s_0, 0)$, and by the implicit function theorem, there exists solutions $s = S(x, y)$, $t = T(x, y)$ of

$$x = X(S(x, y), T(x, y)), \quad y = Y(S(x, y), T(x, y)) \quad (2.11)$$

of class C^1 in a neighborhood of (x_0, y_0) and satisfying $s_0 = S(x_0, y_0)$, $0 = T(x_0, y_0)$, provided the Jacobian determinant

$$\begin{vmatrix} X_s(s_0, 0) & Y_s(s_0, 0) \\ X_t(s_0, 0) & Y_t(s_0, 0) \end{vmatrix} \neq 0, \quad (2.12)$$

which by (2.8), (2.9) is equivalent to

$$\begin{vmatrix} f_s(s_0) & g_s(s_0) \\ a(x_0, y_0, z_0) & b(x_0, y_0, z_0) \end{vmatrix} \neq 0. \quad (2.13)$$

The above gives the local existence condition for the solution of the Cauchy problem for the quasi-linear equation. Uniqueness follows from the following theorem [8].

Theorem 2.2.1 *Let $P = (x_0, y_0, z_0)$ lie on the integral surface $z = u(x, y)$, and Γ be the characteristic curve through P . Then Γ lies completely on S .*

Next we develop the method for the general first-order equation (2.1) in n independent variables.

2.3 Characteristic Equations for the General First-Order Equation

We now consider the general nonlinear first-order PDE (2.1) written in vectorial notation as

$$F(Du, u, x) = 0, \quad x \in U, \quad \text{subject to the boundary condition } u = g \text{ on } \mathcal{O} \quad (2.14)$$

where $Du = (u_{x_1}, u_{x_2}, \dots, u_{x_n})$, $\mathcal{O} \subseteq \partial U$, $g : \mathcal{O} \rightarrow \mathfrak{R}$, and $F, g \in C^\infty(\mathfrak{R})$.

Now suppose u solves (2.14), and fix any point $x \in U$. We wish to calculate $u(x)$ by finding some curve lying within U , connecting x with a point $x^0 \in \mathcal{O}$ and along which we can compute u . Since $u(x^0) = g(x^0)$, we hope to be able to find u along the curve connecting x^0 and x .

To find the characteristic curve, let us suppose that it is described parametrically by the function $\mathbf{x}(s) = (x^1(s), x^2(s), \dots, x^n(s))$, the parameter s lying in some subinterval of \mathfrak{R} . Assume u is a C^2 solution of (2.14), and let

$$z(s) = u(\mathbf{x}(s)), \quad (2.15)$$

$$\mathbf{p}(s) = Du(\mathbf{x}(s)); \quad (2.16)$$

i.e., $\mathbf{p}(s) = (p^1(s), p^2(s), \dots, p^n(s)) = (u_{x_1}(s), u_{x_2}(s), \dots, u_{x_n}(s))$. Now

$$\dot{p}^i(s) = \sum_{j=1}^n u_{x_i x_j}(\mathbf{x}(s)) \dot{x}^j(s), \quad (2.17)$$

where the differentiation is wrt s . On the other hand, differentiating (2.14) wrt x_i , we get

$$\sum_{j=1}^n \frac{\partial F}{\partial p_j}(Du, u, x) u_{x_j x_i} + \frac{\partial F}{\partial z}(Du, u, x) u_{x_i} + \frac{\partial F}{\partial x_i}(Du, u, x) = 0. \quad (2.18)$$

Now if we set

$$\frac{dx_j}{ds}(s) = \frac{\partial F}{\partial p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s)), \quad j = 1, \dots, n, \quad (2.19)$$

and assuming that the above relation holds, then evaluating (2.18) at $x = \mathbf{x}(s)$, we obtain the identity

$$\sum_{j=1}^n \frac{\partial F}{\partial p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s)) u_{x_j x_i} + \frac{\partial F}{\partial z}(\mathbf{p}(s), z(s), \mathbf{x}(s)) \dot{p}^i(s) + \frac{\partial F}{\partial x_i}(\mathbf{p}(s), z(s), \mathbf{x}(s)) = 0. \quad (2.20)$$

Next substituting (2.19) in (2.17) and using the above identity (2.20), we get

$$\dot{p}^i(s) = -\frac{\partial F}{\partial x_i}(\mathbf{p}(s), z(s), \mathbf{x}(s)) - \frac{\partial F}{\partial z}(\mathbf{p}(s), z(s), \mathbf{x}(s)) \dot{p}^i(s), \quad i = 1, \dots, n. \quad (2.21)$$

Finally, differentiating z we have

$$\dot{z}(s) = \sum_{j=1}^n \frac{\partial u}{\partial x_j}(\mathbf{x}(s)) \dot{x}^j(s) = \sum_{j=1}^n p^j(s) \frac{\partial F}{\partial p_j}(\mathbf{p}(s), z(s), \mathbf{x}(s)). \quad (2.22)$$

Thus, we finally have the following system of ODEs

$$\left. \begin{aligned} \dot{\mathbf{p}}(s) &= -D_x F(\mathbf{p}(s), z(s), \mathbf{x}(s)) - D_z F(\mathbf{p}(s), z(s), \mathbf{x}(s)) \mathbf{p}(s) \\ \dot{z}(s) &= D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) \\ \dot{\mathbf{x}}(s) &= D_p F(\mathbf{p}(s), z(s), \mathbf{x}(s)). \end{aligned} \right\} \quad (2.23)$$

where D_x, D_p, D_z are the derivatives wrt x, p, z respectively. The above system of $2n + 1$ first-order ODEs comprises the *characteristic equations* of the nonlinear PDE (2.14). The functions $\mathbf{p}(s), z(s), \mathbf{x}(s)$ together are called the *characteristics* while $\mathbf{x}(s)$ is called the *projected characteristic* onto the physical region $U \subset M$. Furthermore, if $u \in C^2$ solves the nonlinear PDE (2.14) in U and assume \mathbf{x} solves the last equation in (2.23), then $\mathbf{p}(s)$ solves the first equation and $z(s)$ solves the second for those s such that $\mathbf{x}(s) \in U$.

2.3.1 Characteristic Equations for the Hamilton-Jacobi Equation

Let us now consider the characteristic equation for our Hamilton-Jacobi equation from classical mechanics which is a typical nonlinear first-order PDE:

$$G(Du, u_t, u, x, t) = u_t + H(Du, x) = 0, \quad (2.24)$$

where $Du = D_x u$ and the remaining variables have their usual meaning. For simplicity, let $q = (Du, u_t) = (p, p_{n+1})$, $y = (x, t)$. Therefore,

$$G(q, z, y) = p_{n+1} + H(p, x); \quad (2.25)$$

and

$$D_q G = (D_p H(p, x), 1), \quad D_y G = (D_x H(p, x), 0), \quad D_z G = 0.$$

Thus, the characteristic equations (2.23) become

$$\left. \begin{aligned} \dot{x}^i(s) &= \frac{\partial H}{\partial p_i}(\mathbf{p}(s), \mathbf{x}(s)), \quad (i = 1, 2, \dots, n), \\ \dot{x}^{n+1}(s) &= 1, \\ \dot{p}^i(s) &= -\frac{\partial H}{\partial x_i}(\mathbf{p}(s), \mathbf{x}(s)), \quad (i = 1, 2, \dots, n), \\ \dot{p}^{n+1}(s) &= 0 \\ \dot{z}(s) &= D_p H(\mathbf{p}(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) + p^{n+1}. \end{aligned} \right\} \quad (2.26)$$

which can be rewritten in vectorial form as

$$\left. \begin{aligned} \dot{\mathbf{p}}(s) &= -D_x H(\mathbf{p}(s), \mathbf{x}(s)) \\ \dot{\mathbf{x}}(s) &= D_p H(\mathbf{p}(s), \mathbf{x}(s)) \\ \dot{z}(s) &= D_p H(\mathbf{p}(s), \mathbf{x}(s)) \cdot \mathbf{p}(s) - H(\mathbf{p}(s), \mathbf{x}(s)). \end{aligned} \right\} \quad (2.27)$$

The first two of the above equations are clearly Hamilton's canonical equations, while the third equation is the value-function or *action variable* $z(t) = \int_0^t L((\mathbf{x}(s), \dot{\mathbf{x}}(s)))ds$ of the variational problem corresponding to the Hamilton-Jacobi equation, where $L(\mathbf{x}, \dot{\mathbf{x}})$ is the Lagrangian function

$$L(\mathbf{x}, \dot{\mathbf{x}}) = \mathbf{p}\dot{\mathbf{x}} - H(\mathbf{x}, \mathbf{p}).$$

Thus, we have made a connection between Hamilton's canonical equations and the Hamilton-Jacobi equation, and it is clear that a solution for one implies a solution for the other; nevertheless, neither is easy to solve in general. Although, for some systems, the PDE does sometimes offer some leeway, and in fact, this is the motivation behind Hamilton-Jacobi theory.

Next, we specialize the method for the first-order equation in two variables.

2.4 Characteristics for the General First-Order Equation in Two Variables

We now consider the general first-order equation (2.14) in two independent variables of the form

$$F(x, y, z, p, q) = 0, \quad (2.28)$$

where $z = u(x, y)$, $p = u_x$, $q = u_y$. We assume that the function F is C^2 wrt x, y, z, p, Q . From (2.23), we can readily write the characteristic equations for this equation as the following $2n + 1 = 5$ ODEs

$$\left. \begin{aligned} \frac{dx}{dt} &= F_p(x, y, z, p, q) \\ \frac{dy}{dt} &= F_q(x, y, z, p, q) \\ \frac{dz}{dt} &= pF_p(x, y, z, p, q) + qF_q(x, y, z, p, q) \\ \frac{dp}{dt} &= -F_x(x, y, z, p, q) - pF_z(x, y, z, p, q) \\ \frac{dq}{dt} &= -F_y(x, y, z, p, q) - qF_z(x, y, z, p, q) \end{aligned} \right\}. \quad (2.29)$$

It can be checked that F is constant along any trajectory of the system, since

$$\begin{aligned}\frac{dF}{dt} &= F_x \dot{x} + F_y \dot{y} + F_z \dot{z} + F_p \dot{p} + F_q \dot{q} \\ &= F_x F_p + F_y F_q + F_z (p F_p + q F_q) + F_p (-F_x - p F_z) + F_q (-F_y - q F_z) = 0.\end{aligned}$$

Hence, along any trajectory of the system (2.29), $F = 0 \forall t$ if $F = 0$ for some particular t ; implying that F is indeed an integral of the system.

A solution $(x(t), y(t), z(t), p(t), q(t))$ of the characteristic equation (2.29) forms a one-parameter family of elements called a *strip* if the elements, consisting of a point (x, y, z) and a plane through the point with equation

$$\zeta - z = p(\xi - x) + q(\eta - y), \quad (2.30)$$

in running coordinates ξ, η, ζ , are tangent to the curve formed by the points $(x(t), y(t), z(t))$, the *support* of the strip. For this to hold, the *strip condition*

$$\frac{dz}{dt} = p \frac{dx}{dt} + q \frac{dy}{dt}$$

has to be satisfied. Furthermore, a solution of the characteristic equations is called a *characteristic strip*.

Next, we investigate the Cauchy problem for the general first-order equation in two independent variables (2.28). This involves passing an integral surface through an arbitrary initial curve Γ given parametrically by

$$x = f(s), \quad y = g(s), \quad z = h(s). \quad (2.31)$$

Assuming that f, g, h are C^1 for s near a value s_0 corresponding to a point

$$P_0 = (x_0, y_0, z_0) = (f(s_0), g(s_0), h(s_0)). \quad (2.32)$$

To complete Γ into a characteristic strip, we have to find functions

$$p = \phi(s), \quad q = \psi(s) \quad (2.33)$$

such that

$$h'(s) = \phi(s)f'(s) + \psi(s)g'(s) \quad (2.34)$$

$$F(f(s), g(s), h(s), \phi(s), \psi(s)) = 0. \quad (2.35)$$

If we assume that we are given a solution p_0, q_0 of (2.34), (2.35) such that

$$h'(s_0) = p_0 f'(s_0) + q_0 g'(s_0), \quad F(x_0, y_0, z_0, p_0, q_0) = 0 \quad (2.36)$$

and

$$\Delta_2 = f'(s_0)F_q(x_0, y_0, z_0, p_0, q_0) - g'(s_0)F_p(x_0, y_0, z_0, p_0, q_0) = \begin{vmatrix} f' & g' \\ F_p & F_q \end{vmatrix} \neq 0, \quad (2.37)$$

then by the implicit function theorem, there exist unique functions ϕ, ψ of class C^1 near s_0 satisfying (2.34), (2.35) and equal to p_0, q_0 at s_0 . We can then find five functions

$$x = X(s, t), \quad y = Y(s, t), \quad z = Z(s, t), \quad p = P(s, t), \quad q = Q(s, t) \quad (2.38)$$

defined for $|s - s_0|$ and $|t|$ sufficiently small which satisfy for fixed s the characteristic equations (2.29) as functions of t , and for $s = 0$ reduce respectively to $f(s), g(s), h(s), \phi(s), \psi(s)$ an element through which the characteristic strip is passed.

To summarize, we have succeeded in constructing a characteristic strip passing through the curve γ containing the element $(x_0, y_0, z_0, p_0, q_0)$, and by this process an integral surface S can be constructed as the union of the support of such characteristic strips. Conversely, it can be shown (2.38) represents a solution of the Cauchy problem in parametric form in a neighborhood of the point P_0 provided the condition (2.37) is not violated [8].

The above completes the local existence proof for the solution of the Cauchy problem for the general first-order equation in two independent variables and under the assumption that we have a solution p_0, q_0 of (2.36), (2.37). The same analysis can be extended to the general first-order equation on n -independent variables (2.14). In this case, the Cauchy problem consists in finding an integral surface in $x_1 \dots x_n$ x -space that passes through an $n - 1$ -dimensional manifold Γ given parametrically by

$$z = h(s_1, \dots, s_n), \quad x_i = f_i(s_1, \dots, s_{n-1}), \quad i = 1, \dots, n, \quad (2.39)$$

which can be achieved as before by passing a characteristic strip through each point P on the curve γ tangent to γ at P . Thus, we have to find n functions $p_i = \phi(s_1, \dots, s_{n-1})$ such that

$$\frac{\partial h}{\partial s_i} = \sum_{k=1}^n \phi_k \frac{\partial f}{\partial s_i}, \quad i = 1, \dots, n-1, \quad (2.40)$$

$$F(f_1, \dots, f_n, h, \phi_1, \phi_2, \dots, \phi_n) = 0. \quad (2.41)$$

Consequently, in analogy with (2.34)-(2.37), the existence condition for the solution of (2.40), (2.41), is now

$$\Delta_n = \begin{vmatrix} \frac{\partial f_1}{\partial s_1} & \dots & \frac{\partial f_n}{\partial s_1} \\ \vdots & & \vdots \\ \frac{\partial f_1}{\partial s_{n-1}} & \dots & \frac{\partial f_n}{\partial s_{n-1}} \\ F_{p_1} & & F_{p_n} \end{vmatrix} \neq 0. \quad (2.42)$$

2.5 Geometric View-Point of the General First-Order Equation

In the previous section, we have developed the method of characteristic for the general first-order equation in n variables and specialized it to the case of two variables. It is apparent that the method is inherently geometric, and involves the construction of a characteristic strips through integral curves of the system in $n - 1$ dimensions and patching them up

together to form the integral surface. In this section, we develop a more rigorous geometric interpretation of the method using concepts from differential geometry. The presentation here is largely from the unpublished notes [15]

Let \mathcal{D} be an $(m-1)$ -dimensional distribution on an m -manifold M , i.e., $\mathcal{D}(x) \subset T_x M$, $\forall x \in M$, and the codimension of \mathcal{D} is 1. Locally, \mathcal{D} is described by a 1-form $\alpha \in \Omega^1(M)$, which is an element of the codistribution $\mathcal{D}^\perp = \{\beta \in \Omega^1(M) \mid \beta(v) = 0, \forall v \in \mathcal{D}\}$.

Recall that k -dimensional submanifold V of M is an integral manifold of \mathcal{D} if $TV_x \subset \mathcal{D}_x$, $\forall x \in V$. When $k = m - 1$, the dimension of the integral manifold and the distribution are the same, and the distribution is said to be completely integrable. The characterization of complete integrability is given by Frobenius Theorem: \mathcal{D} is completely integrable iff $d\alpha = 0$ on \mathcal{D} .

At the other extreme is nondegeneracy; the distribution \mathcal{D} is nondegenerate if for every $x \in M$, $d\alpha|_{\mathcal{D}_x}$ is nondegenerate symplectic form.

Definition 2.5.1 *Suppose \mathcal{D} is a nondegenerate codimension 1 distribution on a manifold M , then (M, \mathcal{D}) is a contact manifold and \mathcal{D} is a contact structure on M .*

Remark 2.5.1 *Every contact manifold is of odd dimension, and if $m = 2n + 1$, then the nullity of $d\alpha$ is n -dimensional. Therefore, the maximal dimension of an integral manifold of \mathcal{D} is n . Such submanifolds are called Legendre submanifolds.*

We now consider the general first order equation:

$$F(x, u, Du) = 0, \quad (2.43)$$

where $F : U \rightarrow \mathbb{R}$, is $C^r(U)$ and nonsingular, $U \subset \mathbb{R}^{2n+1}$, and we seek an integral surface $S : z = u(x)$ of the PDE for $x \in U$.

Theorem 2.5.1 *Let $p = Du$, $E = \{(x, u, p) \in U \mid F(x, u, p) = 0\}$ be an integral manifold of (2.43), and define $\alpha = du - \sum_{i=1}^n p_i dx_i \in \Omega^1(U)$. Suppose $V^n \subset \mathbb{R}^{2n+1}$ is an embedded submanifold of \mathbb{R}^{2n+1} such that*

- (i) $\pi : V \rightarrow \mathbb{R}^n$, $(x, u, p) \mapsto x$ is a diffeomorphism onto an open set;
- (ii) V is an integral manifold for the distribution defined by α ;
- (iii) $V \subset E$.

If $\pi^{-1}(x) \mapsto (x, u, p)$ is the inverse of the map in (i), then u is a solution of (2.43) on V .

Proof: (i) implies that there exists functions μ, κ such that $(x, u, p) = (x, \mu(x), \kappa(x))$ $\forall (x, u, p) \in V$. Therefore, $du|_{TV} = \sum_{i=1}^n \frac{\partial \mu}{\partial x_i} dx_i|_{TV}$, and

$$\begin{aligned} d\alpha|_{TV} &= du|_{TV} - \sum_i^n p_i dx_i|_{TV} \\ &= \sum_{i=1}^n \frac{\partial \mu}{\partial x_i} dx_i|_{TV} - \sum_i^n p_i dx_i|_{TV}. \end{aligned}$$

By (ii), $\alpha|_{TV} = 0, \Rightarrow \sum_{i=1}^n (\frac{\partial \mu}{\partial x_i} - p_i) dx_i|_{TV} = 0$, and by the linear independence of $\{dx_i | i = 1, \dots, n\}$ on TV , we have $u = \mu(x)$, $p_i = \frac{\partial \mu}{\partial x_i}$ as desired. Furthermore, property (iii) implies that $F(x, u, p) = 0$ on V , so that $F(x, \mu(x), \frac{\partial \mu}{\partial x}) = 0$. \square

Remark 2.5.2 *The above theorem gives a characterization of an integral manifold V of the PDE (2.43) as an n -dimensional embedded submanifold whose projection $\pi : V \rightarrow \mathbb{R}^n$ onto the base manifold is a diffeomorphism.*

2.5.1 The Characteristic Curves of the PDE

In order to solve the PDE (2.43), we need to construct the n -manifold V . The initial conditions form an $(n-1)$ -manifold, and so we need only one more dimension which can be determined by a vector-field or a set of curves through the initial manifold.

The construction of the above vector-field relies on the contact structure on the manifold. If M is a manifold with a contact structure \mathcal{D} , or equivalently by a 1-form $\alpha \in \Omega^1(M)$, and if we let $\omega = d\alpha$, then there is a one-to-one correspondence between differential 1-forms restricted to \mathcal{D} and vector-fields in \mathcal{D} . This correspondence is given in the usual manner for nondegenerate 2-forms by

$$\gamma \leftrightarrow \omega(\gamma, -).$$

We now construct the vector-field on E . Let α be a contact structure on E . Then for each $(x, u, p) \in E$, $\exists! \xi_{(x,u,p)} \in \mathcal{D}_{(x,u,p)}$, a vector-field, and such that $dF(x, u, p) = d\alpha(\xi_{(x,u,p)}, -)$ $\forall (x, u, p) \in E$.

Proposition 2.5.1 *For α described above, $\exists! \xi_{(x,u,p)} \in \mathcal{D}_{(x,u,p)} \cap TE_{(x,u,p)}$, a vector-field, such that $dF(x, u, p) = d\alpha(\xi_{(x,u,p)}, -) = 0$.*

Proof: By construction $\xi \in \mathcal{D}$. We see that the kernel of $d\alpha(\xi_{(x,u,p)}, -)$ on \mathcal{D} is $TE \cap \mathcal{D}$ since the form is dF and the kernel of dF on $T\mathbb{R}^{2n+1}$ is TE . Note that $\xi \in TE$ since $0 = d\alpha(\xi, \xi) = dF(\xi)$. Thus, $\xi_{(x,u,p)} \in \mathcal{D}_{(x,u,p)} \cap TE_{(x,u,p)}, \forall (x, u, p) \in E$; moreover on this intersection $d\alpha(\xi_{(x,u,p)}, -) = 0$. \square

Proposition 2.5.2 *If g^t is the flow of the vector-field ξ , then g_*^t preserves the intersection $TE \cap \mathcal{D}$.*

Proof: Since $\xi \in TE$, TE is preserved. Moreover to show that $TE \cap \mathcal{D}$ is preserved, we need to show that if $v \in TE$ and $\alpha(v) = 0$, then $\alpha(g_*^t(v)) = 0$ or $g^{t*}(\alpha)(v) = 0$. We show that $g_*^t(\alpha) = \alpha$ on TE . Note that $g_*^t(\alpha) = \alpha \forall t$ iff $\frac{\partial d}{\partial t} g^{t*}(\alpha) = 0 \forall t$. Recall further the definition of the Lie derivative $\mathcal{L}_\xi \alpha = -\frac{\partial d}{\partial t} g^{t*}(\alpha) = 0$, and

$$\begin{aligned} \mathcal{L}_\xi \alpha &= di_\xi \alpha + i_\xi d\alpha \\ &= d(\alpha(\xi)) + d\alpha(\xi, -) \\ &= 0 + dF, \end{aligned}$$

where i_- is the inner product operator. But dF vanishes on TE , hence the result. \square

Next we show that the equations to integrate ξ are the characteristic equations discussed in the previous two section.

Let $\xi = (\dot{\mathbf{x}}, \dot{z}, \dot{\mathbf{p}})$. If

$$\eta = \left(\sum_{i=1}^n a_i \frac{\partial}{\partial x_i} + b \frac{\partial}{\partial z} + \sum_{i=1}^n c_i \frac{\partial}{\partial p_i} \right)_{(x,u,p)}$$

for some smooth functions $a_i, b, c_i : U \rightarrow \mathfrak{R}$, and $\alpha = dz - \sum_{i=1}^n p_i dx_i$, then $d\alpha(\xi, \eta) = \sum_{i=1}^n \dot{c}_i \dot{\mathbf{x}}_i - \dot{\mathbf{p}}_i a_i$ and so

$$dF = d\alpha(\xi, -) = \sum_{i=1}^n \dot{\mathbf{x}}_i dp_i - \dot{\mathbf{p}}_i dx_i \quad (2.44)$$

Also,

$$\begin{aligned} dF &= \sum_{i=1}^n \frac{\partial F}{\partial x_i} dx_i + \frac{\partial F}{\partial z} dz + \sum_{i=1}^n \frac{\partial F}{\partial p_i} dp_i \\ &= \sum_{i=1}^n \left(\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial z} dp_i \right) dx_i + \sum_{i=1}^n \frac{\partial F}{\partial p_i} dp_i \end{aligned} \quad (2.45)$$

on \mathcal{D} , since $\alpha = 0$, $\Rightarrow dz = \sum_{i=1}^n p_i dx_i$. Equating (2.44), (2.45), we obtain

$$-\dot{\mathbf{p}}_i = \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial u} p_i, \quad (2.46)$$

$$\dot{\mathbf{x}}_i = \frac{\partial F}{\partial p_i} \quad (2.47)$$

The last equation is more direct. Since $\alpha(\dot{\mathbf{x}}, \dot{z}, \dot{\mathbf{p}}) = 0$ implies that

$$\dot{z} = \sum_{i=1}^n p_i \dot{\mathbf{x}} = \sum_{i=1}^n p_i \frac{\partial F}{\partial p_i}. \quad (2.48)$$

Clearly equations (2.46), (2.47) and (2.48) are the characteristic equations (2.23).

2.5.2 Initial Conditions

The initial conditions form a submanifold $I \subset \mathfrak{R}^{2n+1}$. We have the following theorem regarding the existence of local solutions to the equation (2.43).

Theorem 2.5.2 *Suppose $I \subset \mathfrak{R}^{2n+1}$ is an $(n-1)$ -dimensional submanifold, and*

- (a) $I \subset E$
- (b) $\alpha|_{TI} = 0$
- (c) $\pi|_I : I \rightarrow \mathfrak{R}^n$ is a diffeomorphism to its image
- (d) the vector space generated by $\pi_* T_{(x,u,p)} I$ and $\pi_*(\xi_{(x,u,p)})$ is $T_{(x,u,p)} \mathfrak{R}^n$ for each $(x, u, p) \in I$,

then there exists a Ξ^* an open subset of $\pi(I)$ and \mathcal{D}^* an open neighborhood of Ξ^* such that there exists a unique solution $z = u(x)$ of the PDE (2.43).

Proof: The manifold V is $g^t(I)$. It satisfies conditions (ii), (iii) of Theorem 2.5.1 by Proposition 2.5.2 and items (a), (b) above. Condition (i) in Theorem 2.5.1 follows from item (d) above, namely, transversality in the x direction. \square

Now if initial condition manifold, $I^{(n-1)}$, is parametrized as $\Xi \rightarrow I$ by $s \mapsto (f(s), h(s), g(s))$, then we have the following: items (a) and (b) of Theorem 2.5.2 imply $p = \frac{\partial u}{\partial x}$ so that $g(s) = u_x(f(s))$. Also, (b) implies $du = \sum_{i=1}^n p_i dx_i$ which implies $\frac{\partial h}{\partial s} = g(s) \frac{\partial f}{\partial s}$. Hence

$$h_s(s) = u_x(f(s)) f_s(s)$$

which is equation (2.34) above. While item (a) implies

$$F(f(s), h(s), u_x(f(s))) = 0$$

which is equation (2.35). Further, item (c) implies that

$$D_s f(s) \text{ has maximal rank,}$$

and lastly, item (d) implies that $\text{Det}(D_s f(s), \pi_*(\xi)) \neq 0$. Since $\pi_*(\xi) = \dot{\mathbf{x}} = \frac{\partial F}{\partial p}$, we get that

$$\text{Det}(D_s f(s), F_p) \neq 0$$

which is the condition (2.37) or (2.42). This completes the geometric analysis of the general first-order equation (2.43).

2.5.3 Legendre Transform and Hopf-Lax Formula

Though the method of characteristics provides a remarking way of integrating the HJE, in general the characteristic equations and in particular the Hamilton's canonical equations (2.27) are very difficult to integrate. Thus, other approaches for integrating the HJE had to be sought. One method due to Hopf and Lax [7] which applies to Hamiltonians that are independent of q deserves mention. For simplicity we shall assume that M is an open subset of \mathbb{R}^n , and consider the initial-value problem for the HJE

$$\left. \begin{aligned} u_t + H(Du) &= 0 & \text{in } \mathbb{R}^n \times (0, \infty) \\ u &= g & \text{on } \mathbb{R}^n \times \{t = 0\} \end{aligned} \right\} \quad (2.49)$$

where $g : \mathbb{R}^n \rightarrow \mathbb{R}$ and $H : \mathbb{R}^n \rightarrow \mathbb{R}$ is the Hamiltonian function which is independent of q . Let the Lagrangian function $L : TM \rightarrow \mathbb{R}$ satisfy the following assumptions.

Assumption 2.5.1 *The lagrangian function is convex and $\lim_{q \rightarrow \infty} \frac{L(q)}{|q|} = +\infty$.*

Note that the convexity in the above assumption also implies continuity. Furthermore, for simplicity, we have dropped the \dot{q} dependence of L , and we shall further assume that M is an open subset of \mathbb{R}^n . We then have the following definition.

Definition 2.5.2 *The Legendre transform of L is the Hamiltonian function H defined by*

$$H(p) = \sup_{q \in \mathbb{R}^n} \{p \cdot q - L(q)\}, \quad p \in T_q^* \mathbb{R}^n = \mathbb{R}^n \quad (2.50)$$

$$= p \cdot q^* - L(q^*) \quad (2.51)$$

$$= p \cdot \mathbf{q}(p) - L(\mathbf{q}(p)). \quad (2.52)$$

for some $q^* = \mathbf{q}(p)$.

We note that the “sup” in the above definition is really a “max” ; i.e., there exists some $q^* \in \mathbb{R}^n$ for which the mapping $q \mapsto p \cdot q - L(q)$ has a maximum at $q = q^*$. Further, if L is differentiable, then the equation $p = DL(q^*)$ is solvable for q in terms of p , i.e., $q^* = \mathbf{q}(p)$, and hence the last expression above.

An important property of the Legendre transform [5] is that, it is *involutive*, i.e., if \mathcal{L}_g is the Legendre tranform, then $\mathcal{L}_g^2(L) = L$ and $\mathcal{L}_g(H) = L$. A stronger result is the following.

Theorem 2.5.3 *(Convex duality of Hamiltonian and Lagrangian). Assume L satisfies the Assumptions 2.5.1 and define H as the Legendre transform of L . Then, H also satisfies the following:*

(i) *the mapping $p \mapsto H(p)$ is convex,*

(ii) $\lim_{|p| \rightarrow \infty} \frac{H(p)}{|p|} = +\infty$.

We now use the variational principle to obtain the solution of the initial-value problem (2.49), namely, the Hopf-Lax formula. Accordingly, consider the following variational problem of minimizing the action function:

$$\int_0^t L(\dot{w}(s))ds + g(w(0)) \quad (2.53)$$

over functions $w : [0, t] \rightarrow \mathbb{R}^n$ with $w(t) = q$. The value-function (or cost-to-go) for this minimization problem is given by

$$u(q, t) := \inf \left\{ \int_0^t L(\dot{w}(s))ds + g(q) \mid w(0) = q', w(t) = q \right\}, \quad (2.54)$$

with the infimum taken over all C^1 functions $w(\cdot)$ with $w(t) = q$. Then we have the following result.

Theorem 2.5.4 *(Hopf-Lax Formula). Assume g is Lipschitz continuous, and if $q \in \mathbb{R}^n$, $t > 0$, then the solution $u = u(q, t)$ to the variational problem (2.53) is*

$$u(q, t) = \min_{q' \in \mathbb{R}^n} \left\{ tL \left(\frac{q - q'}{t} \right) + g(q') \right\}. \quad (2.55)$$

The proof of the above theorem can be found in [7].

The next theorem asserts that the Hopf-Lax formula indeed solves the initial-value problem of the HJE (2.49) whenever S in (2.55) is differentiable.

Theorem 2.5.5 *Assume H is convex, $\lim_{|p| \rightarrow \infty} \frac{H(p)}{|p|} = \infty$. Further, suppose $q \in \mathbb{R}^n$, $t > 0$, and S in (2.55) is differentiable at a point $(q, t) \in \mathbb{R}^n \times (0, \infty)$. Then (2.55) satisfied the HJE (2.49) with the initial value $u(q, 0) = g(q)$.*

Again the proof of the above theorem can be found in [7]. The Hopf-Lax formula provides a reasonably weak solution (a Lipschitz-continuous function which satisfies the PDE almost everywhere) to the initial-value problem for the HJE. The Hopf-Lax formula is useful for variational problems and mechanical systems for which the Hamiltonians are independent of configuration coordinates, but is very limited for the case of more general problems.

CHAPTER 3

Solving the Hamilton-Jacobi Equation and Application to the Toda Lattice

3.1 Introduction

In this chapter, we discuss the transformation approach for solving the HJE for a class of Hamiltonian systems, and then apply this approach to solve the Toda lattice equations. In chapter 1, we have reviewed Hamilton's transformation approach for integrating the equations of motion by introducing a canonical transformation which can be generated by a generating function also known as Hamilton's principle function. This led to the Hamilton-Jacobi PDE which must be solved to obtain the required transformation generating function. However, as has been discussed in the previous chapters, the HJE is very difficult to solved except for the case when the Hamiltonian function is such that the equation is separable. It is therefore our objective in this chapter to present a method for solving the HJE for a class of Hamiltonian systems that may not admit a separation of variables. In particular, we shall apply the method to a class of Hamiltonian systems known as the Toda lattice which are integrable [16].

3.2 Solving the Hamilton-Jacobi Equation

In this section we propose an approach for solving the Hamilton-Jacobi equation for a general class of Hamiltonian systems, and then apply the approach to the Toda lattices as special cases. To present the approach, let the configuration space of the class of Hamiltonian systems be a smooth n -dimensional manifold M with local coordinates $q = (q_1, (q_1, \dots, q_n)$, i.e. if (φ, U) is a coordinate chart, we write $\varphi = q$ and $\dot{q}_i = \frac{\partial}{\partial q_i}$ in the tangent bundle $TM|_U = TU$. Further, let the class of systems under consideration be represented by Hamiltonian functions $H : T^*M \rightarrow \mathfrak{R}$ of the form:

$$H(q, p) = \frac{1}{2} \sum_{i=1}^n p_i^2 + V(q), \quad (3.1)$$

where $(p_1(q), \dots, p_n(q)) \in T_q^*M$, and together with (q_1, \dots, q_n) form the $2n$ symplectic coordinates for the phase-space T^*M of any system in the class, while $V : M \rightarrow \mathfrak{R}_+$ is the potential function which we assume to be nonseparable in the variables $q_i, i = 1, \dots, n$. The time-independent HJE corresponding to the above Hamiltonian function is given by

$$\frac{1}{2} \sum_{i=1}^n \left(\frac{\partial W}{\partial q_i} \right)^2 + V(q) = h, \quad (3.2)$$

where $W : M \rightarrow \mathfrak{R}$ is the Hamilton's characteristic function for the system. While (3.2) appears to depend on the local coordinate q in M , it does not. See Theorem 5.2.4 in [1].

We then have the following theorem concerning the solution of this HJE.

Theorem 3.2.1 *Let M be an open subset of \mathbb{R}^n which is simply connected and let $q = (q_1, \dots, q_n)$ be the coordinates on M . Suppose $\rho, \theta_i : M \rightarrow \mathbb{R}$ for $i = 1, \dots, \lfloor \frac{n+1}{2} \rfloor$; $\theta = (\theta_1, \dots, \theta_{\lfloor \frac{n+1}{2} \rfloor})$; and $\zeta_i : \mathbb{R} \times \mathbb{R}^{\lfloor \frac{n+1}{2} \rfloor} \rightarrow \mathbb{R}$ are C^2 functions such that*

$$\frac{\partial \zeta_i}{\partial q_j}(\rho(q), \theta(q)) = \frac{\partial \zeta_j}{\partial q_i}(\rho(q), \theta(q)), \quad \forall i, j = 1, \dots, n, \quad (3.3)$$

and

$$\frac{1}{2} \sum_{i=1}^n \zeta_i^2(\rho(q), \theta(q)) + V(q) = h \quad (3.4)$$

is solvable for the functions ρ, θ . Let

$$\omega^1 = \sum_{i=1}^n \zeta_i(\rho(q), \theta(q)) dq_i,$$

and suppose C is a path in M from an initial point q_0 to an arbitrary point $q \in M$. Then

(i) ω^1 is closed;

(ii) ω^1 is exact;

(iii) if $W(q) = \int_C \omega^1$, then W satisfies the HJE (3.2).

Proof: (i)

$$d\omega^1 = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial}{\partial q_j} \zeta_i(\rho(q), \theta(q)) dq_j \wedge dq_i,$$

which by (3.3) implies $d\omega^1 = 0$; hence, ω^1 is closed. (ii) Since by (i) ω^1 is closed, by the simple connectedness of M , ω^1 is also exact. (iii) By (ii) ω^1 is exact, therefore the integral $W(q) = \int_C \omega^1$ is independent of the path C . Therefore, W corresponds to a scalar function. Furthermore, $dW = \omega^1$ and $\frac{\partial W}{\partial q_i} = \zeta_i(\rho(q), \theta(q))$, and thus substituting in the HJE (3.2) and if (3.4) holds, then W satisfies the HJE. \square

In the next corollary we shall construct explicitly the functions $\zeta_i, i = 1, \dots, n$ in the above theorem.

Corollary 3.2.1 *Assume the dimension n of the system is 2, and M, ρ, θ are as in the hypotheses of Theorem 3.2.1, and that conditions (3.3), (3.4) are solvable for θ and ρ . Also, define the functions $\zeta_i, i = 1, 2$ postulated in the theorem by $\zeta_1(q) = \rho(q) \cos \theta(q)$, $\zeta_2(q) = \rho(q) \sin \theta(q)$. Then, if*

$$\omega^1 = \sum_{i=1}^2 \zeta_i(\rho(q), \theta(q)) dq_i,$$

$W = \int_C \omega^1$, and $q : [0, 1] \rightarrow M$ is a parametrization of C such that $q(0) = q_0, q(1) = q$, then

(i) W is given by

$$W(q, h) = \gamma \int_0^1 \sqrt{h - V(q(s))} [\cos \theta(q(s)) q_1'(s) + \sin \theta(q(s)) q_2'(s)] ds \quad (3.5)$$

where $\gamma = \pm \sqrt{2}$ and $q_i' = \frac{dq_i(s)}{ds}$.

(ii) W satisfies the HJE (3.2).

Proof: (i) If (3.3) is solvable for the function θ , then substituting the functions $\zeta_i(\rho(q), \theta(q))$, $i = 1, 2$ as defined above in (3.4), we get immediately

$$\rho(q) = \pm \sqrt{2(h - V(q))}.$$

Further, by Theorem 3.2.1, ω^1 given above is exact, and $W = \int_C \omega^1 dq$ is independent of the path C . Therefore, if we parametrize the path C by s , then the above line integral can be performed coordinate-wise with W given by (3.5) and $\gamma = \pm\sqrt{2}$. (ii) follows from Theorem 3.2.1. \square

Remark 3.2.1 *The above corollary constructs one explicit parametrization that may be used. However, because of the number of parameters available in the parametrization are limited, the above parametrization is only suitable for systems with $n = 2$. Other types of parametrizations that are suitable could also be employed.*

If however the dimension n of the system is 3, then the following corollary gives a procedure for solving the HJEs.

Corollary 3.2.2 *Assume the dimension n of the system is 3, and M , ρ , are as in the hypotheses of Theorem 3.2.1. Let $\zeta_i : \mathfrak{R} \times \mathfrak{R} \times \mathfrak{R} \rightarrow \mathfrak{R}$, $i = 1, 2, 3$ be defined by $\zeta_1(q) = \rho(q) \sin \theta(q) \cos \varphi(q)$, $\zeta_2(q) = \rho(q) \sin \theta(q) \sin \varphi(q)$, $\zeta_3(q) = \rho(q) \cos \theta(q)$, and assume (3.3) are solvable for θ and φ , while (3.4) is solvable for ρ . Then, if*

$$\omega^1 = \sum_{i=1}^3 \zeta_i(\rho(q), \theta, \varphi) dq_i,$$

$W = \int_C \omega^1$, and $q : [0, 1] \rightarrow M$ is a parametrization of C such that $q(0) = q_0$, $q(1) = q$, then

(i) W is given by

$$\begin{aligned} W(q, h) = & \gamma \int_0^1 \sqrt{(h - V(q(s)))} \left\{ \sin \theta(q(s)) \cos \varphi(q(s)) q_1'(s) + \right. \\ & \left. \sin \theta(q(s)) \sin \varphi(q(s)) q_2'(s) + \cos \theta(q(s)) q_3'(s) \right\} ds, \end{aligned} \quad (3.6)$$

where $\gamma = \pm\sqrt{2}$.

(ii) W satisfies the HJE (3.2).

Proof: Proof follows along the same lines as Corollary 3.2.1. \square

Remark 3.2.2 *Notice that, the parametrization employed in the above corollary is now of a spherical nature.*

The following theorem gives bounds on the solution W and its derivatives.

Theorem 3.2.2 *Let $N \subset M$ be the region in which the solution W of the HJE given in Corollaries 3.2.1 and 3.2.2. exists. Then if C is a path $q : [0, 1] \rightarrow N$ in N parametrized by $s \in [0, 1]$ such that $q(0) = q_0$, $q(1) = q$ we have the following bounds on the solution and its derivatives:*

- (i) $\|W(q(s), h)\|_\infty \leq |\gamma| \sqrt{h} \|q(s)\|_{\mathcal{L}_1};$
- (ii) $\|\frac{\partial W}{\partial q}\|_2 = |\sqrt{2}\rho(q)/\gamma|;$
- (iii) $\|\frac{\partial W}{\partial q}\|_\infty = |\gamma| \sqrt{h} .$

Proof: (i) From (3.5) or (3.6),

$$\begin{aligned} \|W(q, h)\|_\infty &\leq |\gamma| \sum_{i=1}^n \int_0^1 \sup_{q(s) \in N} \left| \sqrt{(h - V(q))} \right| |q'_i(s) dq_i(s)| \\ &\leq |\gamma| \sqrt{h} \int_0^1 (|q'_1(s) ds| + |q'_2(s) ds| + \dots + |q'_n(s) ds|) \\ &\leq |\gamma| \sqrt{h} \|q(s)\|_{\mathcal{L}_1}. \end{aligned}$$

(ii) Using the definition of $\frac{\partial W}{\partial q_i}$ given in Corollaries 3.2.1, 3.2.2, we have $\|\frac{\partial W}{\partial q_i}\|_2^2 = \sum_{i=1}^n |\frac{\partial W}{\partial q_i}|^2 = |\sqrt{2}\rho(q)/\gamma|^2$, hence the result. (iii) Follows by taking the sup over $q \in M$ of $\frac{\partial W}{\partial q_i}$, $i = 1, \dots, n$. \square

Furthermore, the following proposition gives regularity of the solution.

Proposition 3.2.1 *If the functions ρ , θ_i , $i = 1, \dots, \lfloor \frac{n+1}{2} \rfloor$ in Theorem 3.2.1 and Corollaries 3.2.1, 3.2.2, $n = 1, 2$, or 3 exist and the HJE (3.2) is solvable for W , then if θ_i , $i = 1, \dots, \lfloor \frac{n+1}{2} \rfloor$ are C^1 , then W is C^2 , and consequently if θ_i , $i = 1, \dots, \lfloor \frac{n+1}{2} \rfloor$, are C^r , $r \geq 1$, then W is C^{r+1} .*

Proof: From the expressions (3.5), (3.6) for W , we see that ρ is a smooth function, since V is smooth. Hence, the differentiability of W depends on the differentiability of the θ_i , $i = 1, 2$, or 3. Further, it is clear that, the integration increases the differentiability of W by 1 over that of the θ_i , $i = 1, 2$, or 3. \square

We can combine Corollaries 3.2.1 and 3.2.2 for any n in the following proposition.

Proposition 3.2.2 *Let M be an open subset of \mathbb{R}^n which is simply connected and let q_0 be a fixed point in M . Suppose there exists a C^1 matrix function $\mathcal{R} : \mathbb{R}^l \rightarrow SO(n, \mathbb{R})$ for some smooth vector function $\theta = (\theta_1, \dots, \theta_l)$, $\theta_i : M \rightarrow \mathbb{R}$, $i = 1, \dots, l$, and a C^1 vector function $\varrho(q) = [\rho(q), \dots, \rho(q)]$, $\rho : M \rightarrow \mathbb{R}$, such that the Jacobian matrix*

$$\frac{\partial}{\partial q} \mathcal{R}(\theta(q)) \varrho(q) \text{ is symmetric} \quad (3.7)$$

and

$$\frac{1}{2} \langle \varrho(q) | \varrho(q) \rangle + V(q) = h. \quad (3.8)$$

Let

$$\tilde{\omega}^1 = \sum_{i=1}^n [\mathcal{R}(\theta(q))\varrho(q)]_i dq_i$$

and suppose C is a path from q_0 to an arbitrary point $q \in M$. Then,

(i) $\tilde{\omega}^1$ is closed;

(ii) $\tilde{\omega}^1$ is exact;

(iii) if $\tilde{W}(q) = \int_C \tilde{\omega}^1$, then \tilde{W} satisfies the HJE (3.2).

Proof: (i)

$$d\tilde{\omega}^1 = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial}{\partial q_j} [\mathcal{R}(\theta(q))\varrho(q)]_i dq_j \wedge dq_i$$

which by (3.7) implies that $d\tilde{\omega}^1 = 0$; hence, $\tilde{\omega}^1$ is closed. (ii) Again by simple-connectedness of M , (i) implies (ii). (iii) By (ii) the integral $\tilde{W}(q) = \int_C \tilde{\omega}^1$ is independent of the path, and W corresponds to a scalar function. Moreover, if $dW = \tilde{\omega}^1$ and $\frac{\partial W}{\partial q_i} = [\mathcal{R}(\theta(q))\varrho(q)]_i$, then substituting in the HJE (3.2) and if (3.8) holds, then W satisfies the HJE (3.2). \square

If the the HJE (3.2) is solvable, then the dynamics of the system evolves on the n -dimensional *Lagrangian submanifold* \tilde{N} which is an immersed submanifold of maximal dimension, and can be locally parametrized as the graph of the function W , i.e.,

$$\tilde{N} = \{(q, \frac{\partial W}{\partial q}) \mid q \in N \subset M, W \text{ is a solution of HJE (3.2)}\}.$$

as described in Chapter 1. Moreover, for any other solution W' of the HJE, the volume enclosed by this surface is invariant. This is stated in the following proposition.

Proposition 3.2.3 *Let $N \subset M$ be the region in M where the solution W of the HJE (3.2) exists. Then, for any orientation of M , the volume form of \tilde{N}*

$$\omega^n = \left(\sqrt{1 + \sum_{j=1}^n \left(\frac{\partial W}{\partial q_j} \right)^2} \right) dq_1 \wedge dq_2 \dots \wedge dq_n$$

is given by

$$\omega^n = \left(\sqrt{1 + 2(h - V(q))} \right) dq_1 \wedge dq_2 \dots \wedge dq_n$$

Proof: From the HJE (3.2), we have

$$\sqrt{1 + \sum_{j=1}^n \left(\frac{\partial W}{\partial q_j} \right)^2} = \sqrt{1 + 2(h - V(q))}, \quad \forall q \in N$$

\Updownarrow

$$\omega^n = \left(\sqrt{1 + \sum_{j=1}^n \left(\frac{\partial W}{\partial q_j} \right)^2} \right) dq_1 \wedge \dots \wedge dq_n = \left(\sqrt{1 + 2(h - V(q))} \right) dq_1 \wedge \dots \wedge dq_n \quad \forall q \in N \quad \square$$

We now apply the above ideas to solve the HJE for the two-particle \mathcal{A}_2 Toda lattice. We consider the nonperiodic system described in section 1.6.

3.2.1 Solution of the Hamilton-Jacobi equation for the \mathcal{A}_2 Toda system

Consider the the two-particle nonperiodic Toda system (or \mathcal{A}_2 system) given by the Hamiltonian (1.57):

$$H(q_1, q_2, p_1, p_2) = \frac{1}{2}(p_1^2 + p_2^2) + e^{2(q_1 - q_2)}. \quad (3.9)$$

Then, the HJE corresponding to the system is given by

$$\frac{1}{2} \left\{ \left(\frac{\partial W}{\partial q_1} \right)^2 + \left(\frac{\partial W}{\partial q_2} \right)^2 \right\} + e^{2(q_1 - q_2)} = h_2. \quad (3.10)$$

The following proposition gives the solution of the above HJE corresponding to \mathcal{A}_2 Toda lattice.

Proposition 3.2.4 *Consider the HJE (3.10) corresponding to the \mathcal{A}_2 Toda lattice. Then a solution for the HJE is given by*

$$\begin{aligned} W(q'_1, q'_2, h_2) &= \cos\left(\frac{\pi}{4}\right) \int_1^{q'_1} \rho(q) dq_1 + m \sin\left(\frac{\pi}{4}\right) \int_1^{q'_1} \rho(q) dq_1 \\ &= (1+m) \left\{ \frac{\sqrt{h_2 - e^{-2(b+m-1)}} - \sqrt{h_2} \tanh^{-1} \left[\frac{\sqrt{h_2 - e^{-2(b+m-1)}}}{\sqrt{h_2}} \right]}{m-1} - \right. \\ &\quad \left. \frac{\sqrt{h_2 - e^{-2b-2(m-1)q'_1}} - \sqrt{h_2} \tanh^{-1} \left[\frac{\sqrt{h_2 - e^{-2b-2(m-1)q'_1}}}{\sqrt{h_2}} \right]}{m-1} \right\}, \quad q_1 > q_2 \end{aligned}$$

and

$$\begin{aligned} W(q'_1, q'_2, h_2) &= \cos\left(\frac{\pi}{4}\right) \int_1^{q'_1} \rho(q) dq_1 + m \sin\left(\frac{\pi}{4}\right) \int_1^{q'_1} \rho(q) dq_1 \\ &= (1+m) \left\{ \frac{\sqrt{h_2 - e^{-2(b-m+1)}} - \sqrt{h_2} \tanh^{-1} \left[\frac{\sqrt{h_2 - e^{-2(b-m+1)}}}{\sqrt{h_2}} \right]}{m-1} - \right. \\ &\quad \left. \frac{\sqrt{h_2 - e^{-2b+2(1-m)q'_1}} - \sqrt{h_2} \tanh^{-1} \left[\frac{\sqrt{h_2 - e^{-2b+2(1-m)q'_1}}}{\sqrt{h_2}} \right]}{m-1} \right\}, \quad q_2 > q_1 \end{aligned}$$

Furthermore, a solution for the system equations (1.58) for the \mathcal{A}_2 with the symmetric initial $q_1(0) = -q_2(0)$ and $\dot{q}_1(0) = \dot{q}_2(0) = 0$ is

$$q(t) = -\frac{1}{2} \log \sqrt{h_2} + \frac{1}{2} \log [\cosh 2\sqrt{h_2}(\beta - t)] \quad (3.11)$$

where h_2 is the energy and

$$\beta = \frac{1}{2\sqrt{h_2}} \tanh^{-1} \left(\frac{\dot{q}_1^2(0)}{\sqrt{2h_2}} \right).$$

Proof:

Applying the results of Theorem 3.2.1 we have $\frac{\partial W}{\partial q_1} = \rho(q) \cos \theta(q)$, $\frac{\partial W}{\partial q_2} = \rho(q) \sin \theta(q)$ and substituting in the HJE (3.10) we immediately get

$$\rho(q) = \pm \sqrt{2(h_2 - e^{2(q_1 - q_2)})}$$

and

$$\rho_{q_2}(q) \cos \theta(q) - \theta_{q_2} \rho(q) \sin \theta(q) = \rho_{q_1}(q) \sin \theta(q) + \theta_{q_1} \rho(q) \cos \theta(q). \quad (3.12)$$

The above equation (3.12) is a first-order PDE in θ and can be solved by the method of *characteristics* developed in the previous chapter. However, the geometry of the system allows for a simpler solution. We make the simplifying assumption that θ is a constant function. Consequently, equation (3.12) becomes

$$\rho_{q_2}(q) \cos \theta = \rho_{q_1}(q) \sin \theta \Rightarrow \tan \theta = \frac{\rho_{q_2}(q)}{\rho_{q_1}(q)} = -1 \Rightarrow \theta = -\frac{\pi}{4}.$$

Thus,

$$\begin{aligned} p_1 &= \rho(q) \cos\left(\frac{\pi}{4}\right), \\ p_2 &= -\rho(q) \sin\left(\frac{\pi}{4}\right), \end{aligned}$$

and integrating dW along the straightline path from $(1, -1)$ on the line

$$\mathbb{L} : q_2 = \frac{q'_2 + 1}{q'_1 - 1} q_1 + (1 + \frac{q'_2 + 1}{q'_1 - 1}) \triangleq m q_1 + b$$

to some arbitrary point (q'_1, q'_2) we get

$$\begin{aligned} W(q'_1, q'_2, h_2) &= \cos\left(\frac{\pi}{4}\right) \int_1^{q'_1} \rho(q) dq_1 + m \sin\left(\frac{\pi}{4}\right) \int_1^{q'_1} \rho(q) dq_1 \\ &= (1+m) \left\{ \frac{\sqrt{h_2 - e^{-2(b+m-1)}} - \sqrt{h_2} \tanh^{-1}\left[\frac{\sqrt{h_2 - e^{-2(b+m-1)}}}{\sqrt{h_2}}\right]}{m-1} - \right. \\ &\quad \left. \frac{\sqrt{h_2 - e^{-2b-2(m-1)q'_1}} - \sqrt{h_2} \tanh^{-1}\left[\frac{\sqrt{h_2 - e^{-2b-2(m-1)q'_1}}}{\sqrt{h_2}}\right]}{m-1} \right\}. \end{aligned}$$

Similarly, if we integrate from point $(-1, 1)$ to (q'_1, q'_2) , we get

$$\begin{aligned} W(q'_1, q'_2, h_2) &= \cos\left(\frac{\pi}{4}\right) \int_{-1}^{q'_1} \rho(q) dq_1 + m \sin\left(\frac{\pi}{4}\right) \int_{-1}^{q'_1} \rho(q) dq_1 \\ &= (1+m) \left\{ \frac{\sqrt{h_2 - e^{-2(b-m+1)}} - \sqrt{h_2} \tanh^{-1}\left[\frac{\sqrt{h_2 - e^{-2(b-m+1)}}}{\sqrt{h_2}}\right]}{m-1} - \right. \\ &\quad \left. \frac{\sqrt{h_2 - e^{-2b+2(1-m)q'_1}} - \sqrt{h_2} \tanh^{-1}\left[\frac{\sqrt{h_2 - e^{-2b+2(1-m)q'_1}}}{\sqrt{h_2}}\right]}{m-1} \right\}. \end{aligned}$$

Finally, from (1.11) and (3.9), we can write

$$\begin{aligned}\dot{q}_1 &= p_1 = \rho(q) \cos\left(\frac{\pi}{4}\right), \\ \dot{q}_2 &= p_2 = -\rho(q) \sin\left(\frac{\pi}{4}\right).\end{aligned}$$

Then $\dot{q}_1 + \dot{q}_2 = 0$ which implies that $q_1 + q_2 = k$, a constant, and by our choice of initial conditions, $k = 0$. Now integrating the above equations from $t = 0$ to t we get

$$\begin{aligned}\frac{1}{2\sqrt{h_2}} \tanh^{-1}\left(\frac{\rho(q)}{\sqrt{2h_2}}\right) &= \frac{1}{2\sqrt{h_2}} \tanh^{-1}\left(\frac{\rho(q(0))}{\sqrt{2h_2}}\right) - t, \\ \frac{1}{2\sqrt{h_2}} \tanh^{-1}\left(\frac{\rho(q)}{\sqrt{2h_2}}\right) &= \frac{1}{2\sqrt{h_2}} \tanh^{-1}\left(\frac{\rho(q(0))}{\sqrt{2h_2}}\right) - t.\end{aligned}$$

If we let

$$\beta = \frac{1}{2\sqrt{h_2}} \tanh^{-1}\left(\frac{\rho(q(0))}{\sqrt{2h_2}}\right),$$

then upon simplification we get

$$\begin{aligned}q_1 - q_2 &= \frac{1}{2} \log \left[h_2 \left(1 - \tanh^2 2\sqrt{h_2}(\beta - t) \right) \right] \\ &= \frac{1}{2} \log [h_2 \operatorname{sech}^2 2\sqrt{h_2}(\beta - t)].\end{aligned}\tag{3.13}$$

Since $k = 0$, then $q_1 = -q_2 = -q$, and we get

$$\begin{aligned}q(t) &= -\frac{1}{2} \log \sqrt{h_2} - \frac{1}{2} \log [\operatorname{sech} 2\sqrt{h_2}(\beta - t)] \\ &= -\frac{1}{2} \log \sqrt{h_2} + \frac{1}{2} \log [\cosh 2\sqrt{h_2}(\beta - t)].\end{aligned}$$

Now, from (3.10) and (3.13), (3.13),

$$\rho(q(0)) = \dot{q}_1^2(0) + \dot{q}_2^2(0),$$

and in particular, if $\dot{q}_1(0) = \dot{q}_2(0) = 0$, then $\beta = 0$. Therefore,

$$q(t) = -\frac{1}{2} \log \sqrt{h_2} + \frac{1}{2} \log (\cosh 2\sqrt{h_2}t)\tag{3.14}$$

which is of the form (1.63) with $v = \sqrt{h}$. \square

Next, we consider the general solution to the HJE for the \mathcal{A}_2 -Toda lattice. We try to solve the equation (3.12) under the fact that

$$p_1 + p_2 = \alpha\tag{3.15}$$

a constant, which follows from (1.58). Then, from the proceeding, the above equation implies that

$$\rho(q) \cos \theta(q) + \rho(q) \sin \theta(q) = \alpha.\tag{3.16}$$

Now suppose we seek a solution to (3.12) and (3.16) for $\theta(q)$ such that

$$\frac{\partial\theta(q)}{\partial q_1} = \frac{\partial\theta(q)}{\partial q_2}. \quad (3.17)$$

The above condition is satisfied if

$$\theta(q_1, q_2) = f(q_1 + q_2) \quad (3.18)$$

for some smooth function $f : \Re \rightarrow \Re$ of one variable. Then

$$\frac{\partial\theta(q)}{\partial q_1} = \frac{\partial\theta(q)}{\partial q_2} = f'(q_1 + q_2), \quad (3.19)$$

where $f'(\cdot)$ is the derivative of the function with respect to its argument. Then substituting in (3.12) and using (3.16), we get

$$\rho_{q_2}(q) \cos f(q) - \rho_{q_1}(q) \sin f(q) = \alpha f'(q_1 + q_2) \quad (3.20)$$

which after substituting for $\rho_{q_1}(q)$ and $\rho_{q_2}(q)$ and making the change of variables $x = q_1 + q_2$, $y = q_1 - q_2$ becomes

$$\frac{\sqrt{2}e^{2y}}{\sqrt{h_2 - e^{2y}}}(\cos f(x) + \sin f(x)) = \alpha f'(x). \quad (3.21)$$

The above equation represents a first-order nonlinear ODE in the function $f(\cdot)$ and variable x , and can be integrated in this way

$$\int_0^x \frac{\sqrt{2}e^{2y}}{\sqrt{h_2 - e^{2y}}} dx = \int_0^x \frac{\alpha f'(x)}{(\cos f(x) + \sin f(x))} dx \quad (3.22)$$

to yield

$$f(x) = 2 \tan^{-1} \left[\tanh \left(\frac{\sqrt{2}e^{2y}}{\alpha \sqrt{h_2 - e^{2y}}} \right) + 1 \right]. \quad (3.23)$$

Which implies that

$$\theta(q_1, q_2) = 2 \tan^{-1} \left[\tanh \left(\frac{\sqrt{2}e^{2(q_1 - q_2)}}{\alpha \sqrt{h_2 - e^{2(q_1 - q_2)}}} \right) + 1 \right]. \quad (3.24)$$

We can now obtain W by taking the line integral of $p_1(q) = \rho(q) \cos \theta(q)$ and $p_2 = \rho(q) \sin \theta(q)$ along the straightline path from $(1, -1)$ on the line

$$\mathcal{L} : q_2 = \frac{q'_2 + 1}{q'_1 - 1} q_1 - \left(1 + \frac{q'_2 + 1}{q'_1 - 1}\right) \triangleq m q_1 + b$$

to some arbitrary point (q'_1, q'_2) for $q_1 > q_2$ and from $(-1, 1)$ to (q'_1, q'_2) for $q_2 > q_1$. Hence we have

$$W(q, \alpha, h_2) = \int_{\mathcal{L}} [\rho(q) \cos \theta(q) + m \rho(q) \sin \theta(q)] dq_1. \quad (3.25)$$

Using the half-angle formula, we can write

$$T(q_1) \triangleq \tan \frac{\theta(q_1)}{2} = \tanh \left(\frac{\sqrt{2}e^{2q_1(1-m)-b}}{\alpha\sqrt{h_2 - e^{2q_1(1-m)-b}}} \right) + 1 \quad (3.26)$$

$$\cos \theta(q_1) = \frac{1 - T^2(q_1)}{1 + T^2(q_1)} \quad (3.27)$$

$$\sin \theta(q_1) = \frac{2T(q_1)}{1 + T^2(q_1)}. \quad (3.28)$$

Therefore,

$$\begin{aligned} W(q, \alpha, h_2) &= \int_1^{q'_1} \sqrt{2(h_2 - e^{2x(1-m)-b})} \left(\frac{1 - T^2(x)}{1 + T^2(x)} + m \frac{2T(x)}{1 + T^2(x)} \right) dx \\ &= \int_1^{q'_1} \sqrt{2(h_2 - e^{2x(1-m)-b})} \left(\frac{1 - 2mT(x) - T^2(x)}{1 + T^2(x)} \right) dx \text{ for } q_1 > q_2 \end{aligned}$$

and

$$W(q, \alpha, h_2) = \int_{-1}^{q'_1} \sqrt{2(h_2 - e^{2x(1-m)-b})} \left(\frac{1 - 2mT(x) - T^2(x)}{1 + T^2(x)} \right) dx \text{ for } q_2 > q_1. \quad (3.29)$$

REFERENCES

- [1] Abraham R. and Marsden J., *Foundations of Mechanics*, Benjamin Cummings Pub. Co. USA, 1978.
- [2] Adler M. and Van Moerbeke P., “Completely Integrable Systems, Euclidean Lie Algebras, and Curves”, *Adv. Math.*, 38, 267-317, 1980.
- [3] Adler M. and van Moerbeke P., “Linearization of Hamiltonian Systems, Jacobi Varieties, and Representation Theory”, *Adv. Math.*, 38, 318-379, 1980.
- [4] Adler M. and Van Moerbeke P., “The Toda Lattice, Dynkin diagrams, Singularities, and Abelian Varieties”, *Inventiones Math.*, 103, 223-278, 1991.
- [5] Arnold V. I., *Mathematical Methods of Classical Mechanics*, Springer Verlag, NY, 1985.
- [6] Bogoyavlensky O. I., “On Perturbation of the Periodic Toda Lattice”, *Comm. in Mathematical Physics*, 51, pp. 201-209, 1976.
- [7] Evans L. C., *Partial Differential Equations*, Graduate Text in Maths., AMS, 1998.
- [8] Fritz J., *Partial Differential Equations*, Springer-Verlag Applied Mathematical Sciences Series, 3rd Edition, NY, 1978.
- [9] Goldstein H., *Classical Mechanics*, Addison-Wesley, USA, 1950.
- [10] Groesberg S. W., *Advanced Mechanics*, John-Wiley & Sons, Inc. 1968.
- [11] Guest M. A., *Harmonic Maps, Loop Groups, and Integrable Systems*, London Mathematical Society Text 38, Cambridge University Press, 1997.
- [12] Hagedorn P., *Nonlinear Oscillations*, Oxford Engineering Science Series, 1988.
- [13] Marsden J. and Ratiu T. S., *Introduction to Mechanics and Symmetry*, Springer-Verlag, 1999.
- [14] Mcdaniel A. and Smolinsky L., “The Flow of the \mathcal{G}_2 Periodic Toda Lattice”, *J. of Mathematical Physics*, 38, (2), Feb. 1997.
- [15] Smolinsky L., “A Geometric View-Point of the General First-order Nonlinear PDE, Seminar Notes, Department of Mathematics, LSU.
- [16] Toda M., *Theory of Nonlinear Lattices*, Springer-Verlag, Germany, 1981.

VITA

Mohammad Dikko Aliyu, was born on 3rd September, 1966, in Kano, Nigeria. He graduated with a bachelor of engineering degree in electrical engineering from A. B. U. Zaria, Nigeria, in 1988. He was then a Graduate Assistant with the A. T. B. University in Bauchi, Nigeria, between September, 1989, to April 1991. Subsequently, he joined the Department of Systems Engineering, King Fahd University, Dhahran, Saudi Arabia, in April, 1991 as a Research Assistant, and eventually obtained his master's degree in systems engineering in June, 1994. He was then a lecturer with the same department from September, 1994 to December 1998. He then joined the Department of Electrical and Computer Engineering of the Louisiana State University in June, 1999, where he obtained a master's degree in electrical engineering in May, 2001 and the doctoral degree in May, 2002. He is currently a candidate for the degree of master of Science in mathematics which will be conferred at the May, 2003 commencement.